

10/08/47

40

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :  C12N 15/82, 15/29, C07K 14/415, A01N 63/02, A01H 5/00		A2	(11) International Publication Number: <b>WO 95/19443</b>  (43) International Publication Date: 20 July 1995 (20.07.95)
(21) International Application Number: PCT/IB95/00002  (22) International Filing Date: 3 January 1995 (03.01.95)		(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, JP, KG, KP, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).	
(30) Priority Data: 08/181,271 13 January 1994 (13.01.94) US		Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(60) Parent Application or Grant (63) Related by Continuation US Filed on 08/181,271 (CIP) 13 January 1994 (13.01.94)			
(71) Applicant (for all designated States except US): CIBA-GEIGY AG [CH/CH]; Klybeckstrasse 141, CH-4002 Basle (CH).			
(72) Inventors; and (75) Inventors/Applicants (for US only): RYALS, John, A. [US/US]; 14 Sanderling Court, Durham, NC 27713 (US). ALEXANDER, Danny, C. [US/US]; 217 Midenhall Way, Cary, NC 27513 (US). UKNES, Scott, J. [US/US]; 1003 Pinedale Drive, Apex, NC 27502 (US). WARD, Eric, R. [US/CH]; Säckingerstrasse 20, CH-4058 Basle (CH).			
(74) Common Representative: CIBA-GEIGY AG; Patentabteilung, Klybeckstrasse 141, CH-4002 Basle (CH).			
(54) Title: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC DNA SEQUENCES AND USES THEREOF			
(57) Abstract			
<p>The present invention provides recombinant or chimaeric DNA molecules comprising plant SAR genes, wherein the wild-type genes corresponding to said SAR genes can be chemically induced in a plant in a protein-synthesis independent or in a protein-synthesis dependent manner. Methods for obtaining these DNA molecules are also provided. Additionally, the present invention provides chemically inducible wheat genes, <i>Arabidopsis</i> chitinase IV, Maize PR-1mz, and Maize thaumatin PR-5mz. The anti-pathogenic sequences according to the invention can be genetically engineered and transformed into plants to obtain transgenic plants with enhanced resistance to disease. In one embodiment of the present invention, the chemically regulatable DNA promoter sequence of the <i>Arabidopsis</i> Pr-1 gene is provided. Another aspect of the present invention is a method of improving protection of a plant against a pest comprising transgenically expressing in said plant two or more DNA molecules encoding anti-pathogenic proteins, wherein the transgenically expressed proteins exert a synergistic effect.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

**CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC DNA SEQUENCES**  
**AND USES THEREOF**

The present invention relates to DNA molecules encoding proteins capable of conferring plant disease and/or plant pest resistance. Another aspect of the invention relates to the chemical regulation of gene expression. In particular, this invention relates to non-coding DNA sequences which, in the presence of chemical regulators, regulate the transcription of associated DNA sequences in plants. Both aspects of the invention relate, in part, to genes associated with the response of plants to pathogens.

Advances in recombinant DNA technology coupled with advances in plant transformation and regeneration technology have made it possible to introduce new genetic material into plant cells, plants or plant tissue, thus introducing new traits, e.g., phenotypes, that enhance the value of the plant or plant tissue. Recent demonstrations of genetically engineered plants resistant to pathogens (EP-A 240 332 and EP-A 223 452) or insects (Vaeck, M. *et al.*, *Nature* 328: 33 (1987)) and the production of herbicide tolerant plants (DeBlock, M. *et al.*, *EMBO J.* 6: 2513 (1987)) highlight the potential for crop improvement. The target crops can range from trees and shrubs to ornamental flowers and field crops. Indeed, it is clear that the "crop" can also be a culture of plant tissue grown in a bioreactor as a source for some natural product.

Various methods are known in the art to accomplish the genetic transformation of plants and plant tissues (i.e., the stable introduction of foreign DNA into plants). These include transformation by Agrobacterium species and transformation by direct gene transfer.

A. tumefaciens is the etiologic agent of crown gall, a disease of a wide range of dicotyledons and gymnosperms, that results in the formation of tumors or galls in plant tissue at the site of infection. Agrobacterium, which normally infects the plant at wound sites, carries a large extrachromosomal element called the Ti (tumor-inducing) plasmid.

Ti plasmids contain two regions required for tumorigenicity. One region is the T-DNA (transferred-DNA) which is the DNA sequence that is ultimately found stably transferred to plant genomic DNA. The other region required for tumorigenicity is the vir (virulence) region which has been implicated in the transfer mechanism. Although the vir region is absolutely required for stable transformation, the vir DNA is not actually transferred to the infected plant. Transformation of plant cells mediated by infection with Agrobacterium tumefaciens

- 2 -

and subsequent transfer of the T-DNA alone have been well documented. See, for example, Bevan, M.W. and Chilton, M-D., *Int. Rev. Genet.* 16: 357 (1982).

After several years of intense research in many laboratories, the Agrobacterium system has been developed to permit routine transformation of a variety of plant tissue. Representative species transformed in this manner include tobacco, tomato, sunflower, cotton, rapeseed, potato, soybean, and poplar. While the host range for Ti plasmid transformation using A. tumefaciens as the infecting agent is known to be very large, tobacco has been a host of choice in laboratory experiments because of its ease of manipulation.

Agrobacterium rhizogenes has also been used as a vector for plant transformation. This bacterium, which incites hairy root formation in many dicotyledonous plant species, carries a large extrachromosomal element called an Ri (root-inducing) plasmid which functions in a manner analogous to the Ti plasmid of A. tumefaciens. Transformation using A. rhizogenes has developed analogously to that of A. tumefaciens and has been successfully utilized to transform, for example, alfalfa, Solanum nigrum L., and poplar.

Several so-called direct gene transfer procedures have been developed to transform plants and plant tissues without the use of an Agrobacterium intermediate (see, for example, Koziel *et al.*, *Biotechnology* 11: 194-200 (1993) herein incorporated by reference). In the direct transformation of protoplasts the uptake of exogenous genetic material into a protoplast may be enhanced by use of a chemical agent or electric field. The exogenous material may then be integrated into the nuclear genome. The early work was conducted in the dicot tobacco where it was shown that the foreign DNA was incorporated and transmitted to progeny plants, see e.g. Paszkowski, J. *et al.*, *EMBO J.* 3: 2717 (1984); and Potrykus, I. *et al.*, *Mol. Gen. Genet.* 199: 169 (1985).

Monocot protoplasts have also been transformed by this procedure in, for example, Triticum monococcum, Lolium multiflorum (Italian ryegrass), maize, and Black Mexican sweet corn.

Alternatively exogenous DNA can be introduced into cells or protoplasts by microinjection. A solution of plasmid DNA is injected directly into the cell with a finely pulled glass needle. In this manner alfalfa protoplasts have been transformed by a variety of plasmids, see e.g. Reich, T.J. *et al.*, *Bio/Technology* 4: 1001 (1986).

A more recently developed procedure for direct gene transfer involves bombardment of cells by microprojectiles carrying DNA, see Klein, T.M. *et al.*, *Nature* 327: 70 (1987). In this procedure tungsten particles coated with the exogenous DNA are accelerated toward the target cells, resulting in at least transient expression in the example reported (onion).

Just as there is a variety of methods for the transformation of plant tissue, there is a variety of methods for the regeneration of plants from plant tissue. The particular method of regeneration will depend on the starting plant tissue and the particular plant species to be regenerated. In recent years it has become possible to regenerate many species of plants from callus tissue derived from plant explants. The plants which can be regenerated from callus include monocots, such as corn, rice, barley, wheat and rye, and dicots, such as sunflower, soybean, cotton, rapeseed and tobacco.

Regeneration of plants from tissue transformed with *A. tumefaciens* has been demonstrated for several species of plants. These include sunflower, tomato, white clover, rapeseed, cotton, tobacco, and poplar. The regeneration of alfalfa from tissue transformed with *A. rhizogenes* has also been demonstrated. Plant regeneration from protoplasts is a particularly useful technique, see Evans, D.A. *et al.*, in: "Handbook of Plant Cell Culture", Vol. 1, MacMillan Publ. Co., 1983, p. 124. When a plant species can be regenerated from protoplasts, then direct gene transfer procedures can be utilized, and transformation is not dependent on the use of *A. tumefaciens*. Regeneration of plants from protoplasts has been demonstrated for rice, tobacco, rapeseed, potato, eggplant, cucumber, poplar, and corn.

Various plant tissues may be utilized for transformation with foreign DNA. For instance, cotyledon shoot cultures of tomato have been utilized for *Agrobacterium* mediated transformation and regeneration of plants (see European application EP-A 249432). Further examples include *Brassica* species (see WO 87/07299) and woody plant species, particularly poplar (see U.S. Patent No. 4,795,855, incorporated by reference herein in its entirety).

The technological advances in plant transformation and regeneration technology highlight the potential for crop improvement via genetic engineering. There have been reports of genetically engineered tobacco and tomato plants which are resistant to infections of, for example, tobacco mosaic virus (TMV) and resistant to different classes of herbicides. Insect resistance has been engineered in tobacco and tomato plants.

The potential for genetic engineering is not limited to field crops but includes improvements in ornamentals, forage crops and trees. A less obvious goal for plant biotechnology, which includes both genetic engineering and tissue culture applications, is the enhanced production of a vast array of plant-derived chemical compounds including flavors, fragrances, pigments, natural sweeteners, industrial feedstocks, antimicrobials and pharmaceuticals. In most instances these compounds belong to a rather broad metabolic group, collectively denoted as secondary products. Plants may produce such secondary products to ward off potential predators, attract pollinators, or combat infectious diseases.

Plant cell cultures can be established from an impressive array of plant species and may be propagated in a bioreactor. Typical plant species include most of those that produce secondary products of commercial interest. It has been clearly demonstrated in a number of agriculturally important crop plants that stable genetic variants arising from the tissue culture of plant somatic cells (somatic variation) can be induced and selected. Numerous advantages flow from plant tissue culture production of secondary compounds. These include (1) the possibility of increased purity of the resultant product, (2) the conversion of inexpensive precursors into expensive end products by biotransformation, and (3) the potential for feeding substrate analogs to the culture to create novel compounds.

Whether the target of genetic engineering of plants is a field crop, ornamental shrub, flower, tree or a tissue culture for use in a bioreactor, a principal advantage to be realized is the control of expression of the chimeric gene so that it is expressed only at the appropriate time and to the appropriate extent, and in some situations in particular parts of the plant. For example, in order to achieve a desirable phenotype the chimeric gene may need to be expressed at levels of 1% of the total protein or higher. This may well be the case for fungal resistance due to chimeric chitinase expression or insect resistance due to increased proteinase inhibitor expression. In these cases the energy expended to produce high levels of the foreign protein may result in a detriment to the plant whereas, if the gene were expressed only when desired, for instance when a fungal or insect infestation is imminent, the drain on energy, and therefore yield, could be reduced.

Alternatively, the phenotype expressed by the chimeric gene could result in adverse effects to the plant if expressed at inappropriate times during development. For example, if the chimeric gene product were a plant hormone that induced pod abscission, early expression could bring about abscission of the fruit from the plant before the seed had

- 5 -

matured, resulting in decreased yield. In this case it would be much more advantageous to induce the expression of this type of gene to a time when pod abscission is preferred, or least injurious to the plant.

For tissue in culture or in a bioreactor the untimely production of a secondary product could lead to a decrease in the growth rate of the culture resulting in a decrease in the yield of the product. Therefore, it would be advantageous to allow the culture to grow without expressing the secondary product and then induce the chimeric gene at an appropriate time to allow for an optimized expression of the desired product.

In view of considerations like these, as well as others, it is clear that control of the time, extent and/or site of expression of the chimeric gene in plants or plant tissues would be highly desirable. Control that could be exercised easily in a field, a greenhouse or a bioreactor would be of particular commercial value.

External manipulation of the expression of endogenous genes which contain chemically regulatable sequences by the application of a chemical regulator is already known from Ward, E. et al., *Plant Cell* 3: 1085-1094 (1991); Williams et al., *Bio/Technology* 10: 540-543 (1992); and Uknas, S. et al., *Plant Cell* 5: 159-169 (1993). The control provided is somewhat limited, however, due to the responsiveness of such sequences to endogenous chemical metabolites and cell signals as well as externally applied chemical regulators. In one aspect of the invention, alterations are taught which block the responsiveness of these genes to endogenous signals while maintaining responsiveness to externally applied chemical regulators.

Pest infestation of crop plants causes considerable loss of yield throughout the world and most crops grown in the U.S. suffer infestation, particularly from insects. Major insect pests in the U.S. include the European Corn Borer (*Ostrinia nubilalis*) in maize, the cotton bollworm (*Heliothis zea*) and the pink bollworm (*Pectinophora gossypiella*) in cotton and the tobacco budworm (*Heliothis virescens*) in tobacco. Resistance to pests is difficult to achieve using conventional breeding programs and typically pests have been controlled using chemical pesticides.

Recent advances in molecular biology and plant transformation technology have demonstrated the possibility of expressing in transgenic plants genes encoding insecticidal proteins; this represents a novel approach in the production of crop plants resistant to

pests. Most notably, the expression of genes encoding the *Bacillus thuringiensis* endotoxin has been successful in a wide range of plant species, and the analysis of transgenic lines expressing such genes has been well documented (Vaeck *et al.*, *Nature* 328: 33-37 (1987); Fischhoff *et al.*, *Biotechnology* 5: 807-813 (1987); Carozzi *et al.*, *Plant Mol. Biol.* 20: 539-548 (1992); Koziel *et al.*, *Biotechnology* 11: 194-200 (1993)). Other insecticidal genes have been used successfully in generating insect resistant transgenic plants.

One approach has been the overexpression of genes encoding insect enzyme inhibitors such as trypsin inhibitors or seed proteins with known insecticidal activity, such as lectins (Hilder *et al.*, *Nature* 330: 160-163 (1987)). Indeed, plants expressing both the cowpea trypsin inhibitor and pea lectin were shown to have additive effects in providing insect resistance (Boulter *et al.*, *Crop Protection* 9: 351-354 (1990)). In cases where pests are able infest parts of the plant or tissues not readily accessible to conventional pesticides, a transgenic approach may be more successful than the use of conventional pesticides.

For example, the tobacco budworm *Heliothis* is well known to be difficult to control using pesticides because it burrows deep into the plant tissue. Additionally, some pests of roots, such as nematodes, are not readily controlled by foliar applications of pesticides. An advantage in the use of transgenic plants expressing insecticidal proteins is the controlled expression of the proteins in all desired tissues.

Chitinases catalyze the hydrolysis of chitin, a  $\beta$ -1,4-linked homopolymer of *N*-acetyl-D-glucosamine. Several different plant chitinases have been described and the cDNA sequences for some of these have been reported (Meins *et al.*, *Mol. Gen. Genet.* 232: 460-469 (1992)). Based on structural characteristics, three classes have been distinguished. Class I chitinases have two structural domains, a cysteine-rich amino-terminal hevein domain and a carboxyterminal catalytic domain separated from the former by a variable spacer. Class II chitinases lack the cysteine-rich hevein domain and all or part of the variable spacer, but retain the catalytic domain. Class III chitinases lack the hevein domain and contain a catalytic domain that shares no significant homology with that of the class I or class II enzymes.

Class I chitinase gene expression is induced by ethylene, whereas class II and class III chitinase gene expression is induced in the SAR response. The chitinase/lysozyme disclosed in European patent application EP-A 392 225 and the chitinase/lysozyme disclosed in SEQ ID Nos. 20 and 21 are class III chitinases. It is well known that the level of chitinase activity of plants increases dramatically after pathogen invasion (Mauch *et al.*,

*Plant Physiol.* 76: 607-611 (1984)) and this is presumably due to the host plant's attempts to degrade the chitin of the fungal cell wall. Furthermore, chitinase has been shown *in vitro* to inhibit fungal and insect growth, and in transgenic plants a bacterial chitinase has been shown to exhibit inhibitory effects towards numerous pathogens and pests including insects (Suslow & Jones WPO 90-231246; U.S. Pat. Nos. 4,940,840 and 4,751,081; herein incorporated by reference in their entirety).

For over 90 years, scientists and naturalists have observed that when plants survive pathogen infection they develop an increased resistance to subsequent infections. In 1933, a phenomenon termed "physiological acquired immunity" was described in an extensive literature review by Chester, K.S., *Quart. Rev. Biol.* 8: 275-324 (1933). At that time, scientists believed they were investigating a phenomenon analogous to the immune response in mammals. In retrospect, at least three different processes were being called "acquired immunity": viral cross protection, antagonism (or biocontrol), and what we now refer to as systemic acquired resistance (SAR).

The first systematic study of SAR was published by A. Frank Ross in 1961. Using tobacco mosaic virus (TMV) on local lesion hosts, Ross demonstrated that infections of TMV were restricted by a prior infection. This resistance was effective against not only TMV, but also tobacco necrosis virus and certain bacterial pathogens. Ross coined the term "systemic acquired resistance" to refer to the inducible systemic resistance (Ross, A.F., *Virology* 14: 340-358 (1961)) and "localized acquired resistance" (LAR) to describe the resistance induced in inoculated leaves (Ross, A.F., *Virology* 14: 329-339 (1961)). It is still unclear whether SAR and LAR are aspects of the same response or distinct processes.

In the past 30 years, SAR has been demonstrated in many plant species and the spectrum of resistance has been broadened to include not only viruses and bacteria, but also many agronomically important phytopathogenic fungi (see Kuc, J., *BioScience* 32: 854-860 (1982). However, understanding of the biochemical events leading to the establishment of SAR had not progressed substantially until the past dozen years. In 1982, the accumulation of a group of extracellular proteins called pathogenesis-related (PR) proteins were reported to correlate with the onset of SAR (Van Loon, L.C. *et al.*, *Neth. J. Plant. Path.* 88: 237-256 (1982)). In 1979, salicylic acid (SA) and certain benzoic acid derivatives were reported to be able to induce both resistance and the accumulation of PR proteins (White,

R.F., *Virology* 99: 410-412 (1979). As a result, SA was considered as a possible endogenous signal molecule (Van Loon, L.C. *et al.*, *Neth. J. Plant. Path.* 88: 237-256 (1982)). However, progress slowed through the 1980's and the involvement of PR proteins and salicylic acid in SAR was questioned.

With the advent of genetic engineering and recombinant DNA technology, the possibility of manipulating genetic material to improve the phenotype of plants has arisen. The present invention is based in part upon the discovery of the identity and role of genes involved in SAR which has allowed the application of modern molecular biological techniques for improved plant disease and plant pest resistance.

There are two major aspects of the present invention. The first aspect relates to chemically regulatable DNA sequences and the chemicals which regulate them. The second aspect relates to plant pathogenesis-related proteins. Both aspects of the invention have arisen, in part, from the inventors' identification and characterization of cDNAs and corresponding genes involved in the plant response to pathogen infection.

A principal object of the present invention is to provide additional means for chemically regulating the expression of a desired gene in a plant, seed, or plant tissue. In particular, the present invention relates to identification and isolation of a non-coding DNA sequence obtainable from the promoter region of an *Arabidopsis* PR-1a gene which, in the presence of chemical regulators, regulate the transcription of associated DNA sequences in plants.

Another principal object of the present invention is to provide transgenic plants, transgenic seeds and the progeny thereof expressing levels of plant pathogenesis-related anti-pathogenic proteins such as tobacco chitinase/lysozyme according to the invention, or substantially homologous proteins, which confer an enhanced disease-resistant and/or pest-resistant phenotype with respect to wild-type plants.

Accordingly, to meet this objective and others, the second aspect of the present invention disclosed herein provides for the isolation, cloning and identification of novel cDNA clones coding for plant pathogenesis-related (PR) proteins. These cDNAs, or their genomic counterparts, or DNA molecules with substantial homology to either (all of the above collectively referred to herein as "anti-pathogenic sequences"), can be engineered for expression of the encoded PR proteins or anti-sense RNA and transformed into plants to confer enhanced resistance or tolerance to various diseases and/or pests as described

herein. These DNA molecules may be engineered for constitutive expression, expression in particular tissues or at certain developmental stages, or induced expression in response to an inducer, for example in response to a chemical inducer as described herein.

The present invention further provides novel methods for differential screening and enriching for induced cDNAs, particularly those cDNAs induced in response to pathogen infection or a chemical inducer which triggers a response mimicking pathogen infection.

The present invention teaches the cloning of SAR genes by differential screening of tissues induced and non-induced to the systemic acquired response. SAR induction was found cause the transcription of genes in a protein synthesis-dependent fashion and also a protein synthesis-independent fashion. Two methods were used to clone specifically genes whose induced transcription is protein-synthesis independent:

Firstly, recombinant or chimaeric DNA molecules encoding genes which correspond to wild-type genes which can be chemically induced in a protein-synthesis independent manner, are obtained by

- (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;
- (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying clones which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.

Thus, cDNAs which were cloned by standard differential screening techniques were further screened on SAR-induced RNA isolated with and without cycloheximide (CHX) pre-treatment.

Secondly, recombinant or chimaeric DNA molecules encoding genes which correspond to wild-type genes which can be chemically induced in a protein-synthesis independent manner, are obtained by

- (a) cloning genes which are chemically inducible in the presence or absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;

- 10 -

- (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying amplification products which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.

Thus, a PCR-based "differential display" technique is used to identify SAR-induced, but protein synthesis independent cDNAs directly.

Differential display RNAs were prepared with and without SAR induction and CHX treatment. CHX as an inhibitor of protein synthesis is well known in the art and described by Greenberg *et al.*, *Mol. Cell Biol.* 6: 1050-1057 (1986), Lau and Nathans, *Proc. Natl. Acad. Sci.* 84: 1182-1186 (1987), and Uknas *et al.*, *Plant Cell* 5: 159-169 (1993). The methods described provide recombinant or chimaeric DNA molecules comprising plant SAR genes and preferably plant SAR cDNA molecules, wherein the wild-type genes corresponding to said SAR genes can be chemically induced in a plant in a protein-synthesis independent manner, that is in the presence of a protein synthesis inhibitor such as cycloheximide. Thus, a number of genes were cloned which were induced by the SAR response, yet expressed independently of protein synthesis. These cloned genes are likely signal transducers in the pathway leading from induction to the development of the resistant state.

Additionally recombinant or chimaeric DNA molecules are provided which comprise plant SAR genes and preferably plant SAR cDNA molecules, wherein the wild-type genes corresponding to said SAR genes can be chemically induced in a plant in a protein-synthesis dependent manner, that is only in the absence of a protein synthesis inhibitor such as cycloheximide. Recombinant or chimaeric DNA molecules encoding genes which correspond to wild-type genes which can be chemically induced in a protein-synthesis dependent manner, can be obtained by

- (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;
- (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor such as cycloheximide; and

- 11 -

- (c) identifying clones which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

Alternatively, these molecules can be obtained by

- (a) cloning genes which are chemically inducible in the absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;
- (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor such as cycloheximide; and
- (c) identifying amplification products which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

#### **DESCRIPTION OF GENE SEQUENCES**

- SEQ ID No. 1: The cDNA sequence encoding an *Arabidopsis* class IV chitinase with a hevein domain (pChit4-TA).
- SEQ ID No. 2: The cDNA sequence encoding an *Arabidopsis* class IV chitinase without a hevein domain (pChit4-TB).
- SEQ ID No. 3: The cDNA sequence of the wheat gene WCI-1.
- SEQ ID No. 4: The partial cDNA sequence of the 5' end of the wheat gene WCI-2 which encodes a lipoxygenase isozyme. The partial sequence of the 3' end of this cDNA is provided in SEQ ID No. 5.
- SEQ ID No. 5: The partial cDNA sequence of the 3' end of the wheat gene WCI-2 which encodes a lipoxygenase isozyme. The partial sequence of the 5' end of this cDNA is provided in SEQ ID No. 4.
- SEQ ID No. 6: The cDNA sequence of the wheat gene WCI-3.
- SEQ ID No. 7: The cDNA sequence of the wheat gene WCI-4.
- SEQ ID No. 8: The cDNA sequence of the wheat gene WCI-5.
- SEQ ID No. 9: A tobacco protein-synthesis dependent gene involved in the regulation of the systemic acquired resistance response designated p1.1.1.

SEQ ID No. 10: A tobacco protein-synthesis dependent gene involved in the regulation of the systemic acquired resistance response designated p11.31.4.

SEQ ID No. 11: The 5' DNA sequence of a tobacco protein-synthesis independent gene involved in the regulation of the systemic acquired resistance response designated p11.30.13.

SEQ ID No. 12: The DNA sequence of the 3'end of the same protein-synthesis independent gene described in Seq. ID. No. 11 cloned from tobacco and involved in the regulation of the systemic acquired resistance response designated p11.30.13. This sequence is derived from the non-coding strand (*i.e.* the "bottom" strand). The first base listed is therefore located in the furthest 3' position.

SEQ ID No. 13: A tobacco protein-synthesis independent gene involved in the regulation of the systemic acquired resistance response designated p1.4.3. This sequence is identical to the thioredoxin gene published by Brugidou *et al.*, *Mol. Gen. Genet.* 238: 285-293 (1993).

SEQ ID No. 14: A protein-synthesis independent SAR gene cloned from tobacco designated p66B1.

SEQ ID No. 15: A protein-synthesis dependent SAR gene cloned from tobacco designated p14.22.3.

SEQ ID No. 16: An *Arabidopsis* protein-synthesis independent gene involved in the regulation of the systemic acquired resistance response designated pDPA2.

SEQ ID No. 17: The cDNA sequence encoding an *Arabidopsis* PR-1 protein cloned into plasmid pAPR1C-1.

SEQ ID No. 18: Oligo B used in example 4A

SEQ ID No. 19: Oligo C used in example 4A

SEQ ID No. 20: The cDNA sequence encoding a basic tobacco chitinase/lysozyme protein cloned into the plasmid pBSCL2.

SEQ ID No. 21: The cDNA sequence encoding an acidic tobacco chitinase/lysozyme protein cloned into the plasmid pBSTCL226.

SEQ ID No. 22: A representative molecular adaptor sequence.

SEQ ID No. 23: Oligonucleotide primer for the PR-1 gene.

SEQ ID No. 24: Oligonucleotide primer for GUS gene.

SEQ ID No. 25: Oligonucleotide primer for the AHAS gene.  
SEQ ID No. 26: Oligonucleotide primer for the BT gene.  
SEQ ID No. 27: Peptide sequence used to design a degenerate forward oligonucleotide sequence for PCR cloning of chitinase IV  
SEQ ID No. 28: Peptide sequence used to design a degenerate reverse oligonucleotide sequence for PCR cloning of chitinase IV  
SEQ ID No. 29: The cDNA sequence of the Maize PR1-like gene PR-1mz  
SEQ ID No. 30: Amino acid sequence of the protein encoded by SEQ ID No. 29  
SEQ ID No. 31: The cDNA sequence of the Maize thaumatin-like gene PR-5mz  
SEQ ID No. 32: Amino acid sequence of the protein encoded by SEQ ID No. 31  
SEQ ID No. 33: The cDNA sequence of the Arabidopsis gene PSI-1  
SEQ ID No. 34: The cDNA sequence of the Arabidopsis gene PSI-2  
SEQ ID No. 35: The cDNA sequence of the Arabidopsis gene PSI-3  
SEQ ID No. 36: The cDNA sequence of the Arabidopsis gene PSI-4  
SEQ ID No. 37: The cDNA sequence of the Arabidopsis gene PSI-5

#### **DEFINITIONS**

In order to provide a clear and consistent understanding of the specification and the claims, including the scope given to such terms, the following definitions are provided:

**Anti-pathogenic Sequence:** A DNA molecule encoding a plant pathogenesis-related (PR) protein, or a DNA molecule with substantial homology thereto, which is capable of conferring enhanced resistance or tolerance to disease and/or pests when expressed in a plant, seed, or plant tissue.

**Anti-sense Mechanism:** A mechanism for regulation of gene expression based on the presence in a cell of a RNA molecule complementary to at least a portion of the mRNA encoded by the gene. This mechanism is thought to involve controlling the rate of translation of mRNA to protein due to the presence in a cell of an RNA molecule complementary to at least a portion of the mRNA being translated.

Associated DNA Sequence: A DNA sequence whose cellular activity either (1) regulates the activity of another DNA sequence or (2) is regulated by another DNA sequence. This definition specifically embraces, but is not limited to, sequences which are physically adjacent in a continuous DNA strand or which are physically separated. Physical separation includes, for example, separation within the same DNA strand, location within different DNA strands, or discontinuous interspersed sequences (e.g., alternating regulatable and coding sequences) in one strand.

Chemically Regulatable DNA Sequence: A DNA sequence which is capable of regulating the transcription of an associated DNA sequence where the regulation is dependent on a chemical regulator. The sequences may be of natural or synthetic origin.

Chemically Regulatable Gene: A gene containing at least one non-coding chemically regulatable DNA sequence and at least one associated coding DNA sequence. The genes may be of natural, synthetic or partially natural/partially synthetic origin.

Chemical Regulator (for a chemically regulatable DNA sequence): An elemental or molecular species which controls (e.g., initiates, terminates, increases or reduces), by direct or indirect action, the activity of a chemically regulatable DNA sequence in a system in which the chemical regulator is not normally found in an active form in an amount sufficient to effect regulation of transcription, to the degree and at the time desired, of a transcribable DNA sequence associated with the chemically regulatable DNA sequence. This terminology is intended to embrace situations in which no or very little regulator is present at the time transcription is desired or in which some regulator is present but increased or decreased regulation is required to effect more or less transcription as desired.

Thus, if the system containing the chemically regulatable DNA sequence is a plant, for example a transgenic plant, a chemical regulator is a species not naturally found in the plant in an amount sufficient to effect chemical regulation, and thus transcription of an associated gene, to the desired degree at the time desired.

By "direct action" it is intended that the chemical regulator action result from the direct interaction between the chemical regulator and the DNA sequence. By "indirect action" it is meant that the regulator action results from the direct interaction between the chemical regulator and some other endogenous or exogenous component in the system,

the ultimate result of that direct interaction being activation or suppression of the activity of the DNA sequence. By "active form" it is intended that the chemical regulator be in a form required to effect control.

Chimeric Sequence or Gene: A DNA sequence containing at least two heterologous parts, e.g., parts derived from naturally occurring DNA sequences which are not associated in their naturally occurring states, or containing at least one part that is of synthetic origin and not found in nature.

Coding DNA Sequence: A DNA sequence which, when transcribed and translated, results in the formation of a cellular polypeptide.

Constitutive transcription: Transcription of substantially fixed amounts of a DNA sequence, irrespective of environmental conditions.

Gene: A discrete chromosomal region which is responsible for a discrete cellular product.

Inducers: Molecules that cause the production of larger amounts of macromolecules, compared to the amounts found in the absence of the inducer.

Inducible Protein: Proteins whose rate of production can be increased by the presence of inducers in the environment.

Non-coding DNA Sequence: A DNA sequence, which is not transcribed and translated, resulting in the formation of a cellular polypeptide when associated with a particular coding DNA sequence. A sequence that is non-coding when associated with one coding sequence may actually be coding when associated with another coding or non-coding sequence.

Phenotypic Trait: An observable property resulting from expression of a gene.

Plant Tissue: Any tissue of a plant in planta or in culture. This term includes, but is not limited to, whole plants, plant cells, plant organs, plant seeds, protoplasts, callus, cell cultures and any groups of plant cells organized into structural and/or functional units. The

use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

PR, or Pathogenesis-Related Proteins: Proteins expressed in plants reacting hypersensitively towards pathogens. This term embraces, but is not limited to, SAR8.2a and SAR8.2b proteins, the acidic and basic forms of tobacco PR-1a, PR-1b, PR-1c, PR-1', PR-2, PR-2', PR-2", PR-N, PR-O, PR-O', PR-4, PR-P, PR-Q, PR-S, and PR-R major proteins, cucumber peroxidases, basic cucumber peroxidase, the chitinase which is a basic counterpart of PR-P or PR-Q, and the beta-1,3-glucanase (glucan endo-1,3- $\beta$ -glucosidase, EC 3.2.1.39) which is a basic counterpart of PR-2, PR-N or PR-O, the pathogen-inducible chitinase from cucumber, class IV chitinases with and without a hevein domain, and the WCI ("Wheat Chemically Induced") gene proteins from . A hypersensitive reaction is characterized by a local necrosis of the tissues immediately surrounding the infection site of the pathogen and a subsequent localization of the pathogen, which is in contrast to a sensitive reaction wherein the pathogen spreads throughout the plant. Pathogens are, for example, viruses or viroids, e.g. tobacco or cucumber mosaic virus, ringspot virus or necrosis virus, pelargonium leaf curl virus, red clover mottle virus, tomato bushy stunt virus, and like viruses, fungi, e.g. Phytophthora parasitica or Peronospora tabacina, bacteria, e.g. Pseudomonas syringae or Pseudomonas tabaci, or aphids, e.g. Myzus persicae. This list is not limiting in any respect.

Regulation: The increasing (inducing) or decreasing (repressing) of the level of expression of a gene or the level of transcription of a DNA sequence. The definition is not intended to embrace any particular mechanism.

Substantially Pure DNA Sequence: A DNA molecule (sequence) isolated in substantially pure form from a natural or non-natural source. Such a molecule may occur in a natural system, for example, in bacteria, viruses or in plant or animal cells, or may be provided, for example, by synthetic means or as a cDNA. Substantially pure DNA sequences are typically isolated in the context of a cloning vector. Substantially pure means that DNA sequences other than the ones intended are present only in marginal amounts, for example less than 5%, less than 1%, or preferably less than 0.1%. Substantially pure DNA sequences and

vectors containing may be, and typically are, provided in solution, for example in aqueous solution containing buffers or in the usual culture media.

Substantial Sequence Homology: Substantial sequence homology means close structural relationship between sequences of nucleotides or amino acids. For example, substantially homologous DNA sequences may be 80% homologous, preferably 90% or 95% homologous, and substantially homologous amino acid sequences may typically be 50% homologous, or more. Homology also includes a relationship wherein one or several subsequences of nucleotides or amino acids are missing, or subsequences with additional nucleotides or amino acids are interdispersed.

#### ABBREVIATIONS

The following abbreviations are used herein:

AHAS	acetohydroxyacid synthase
ATCC	American Type Culture Collection
ATP	adenosine triphosphate
bp	base pair
BT	<u>Bacillus thuringiensis</u> endotoxin
BTH	methyl benzo-1,2,3-thiadiazole-7-carboxylate
CAT	chloramphenicol acetyltransferase
CETAB	hexadecyltrimethylammonium bromide
CHX	cycloheximide
2,4-D	2,4-dichlorophenoxyacetic acid
DTT	dithiothreitol
dicamba	3,6-dichloro-2-methoxybenzoic acid
EDTA	ethylenediamine N,N,N',N'-tetraacetic acid
GUS	beta-1,3-glucuronidase
INA	isonicotinic acid
kb	kilo base pair
LUX	luciferase

MES	2-(N-morpholino)ethanesulfonic acid
MU	4-methyl umbelliferyl glucuronide
NOS	nopaline synthase
NPT	neomycin phosphotransferase
NRRL	designation for deposits made with the Agricultural Research Culture Collection, International Depositing Authority, 1815 N. University Street, Peoria, Illinois 61604
OCS	octopine synthase
PEG	polyethylene glycol
picloram	4-amino-3,5,6-trichloropicolinic acid
PR protein	Pathogenesis-related protein
SA	salicylic acid
SAR	Systemic Acquired Resistance
SDS	sodium dodecyl sulfate
TFA	trifluoroacetic acid
TMV	tobacco mosaic virus
Tris-HCl	tris(hydroxymethyl)methylamine hydrochloride
WCI	Wheat Chemically Induced (gene nomenclature designation)

#### ANTI-PATHOGENIC SEQUENCES

The present invention also embraces anti-pathogenic DNA sequences which are capable of conferring enhanced disease resistance or disease tolerance when expressed in a plant or plant tissue. This includes coding sequences for plant pathogenesis-related (PR) proteins as described herein and sequences with substantial homology to these coding sequences.

Included within the scope of the present invention, in addition to the sequences exemplified specifically below and enumerated in the sequence listing, are cDNA sequences which are equivalent to the enumerated sequences which encode the given plant pathogenesis-related protein, and cDNA sequences which hybridize with the enumerated sequences and encode a polypeptide having some degree of disease-resistant activity of the given plant pathogenesis-related protein (i.e. an anti-pathogenic sequence).

Equivalent cDNA sequences are those which encode the same protein even though they contain at least one different nucleotide from the enumerated sequence. As is well known in the art, the amino acid sequence of a protein is determined by the nucleotide sequence of the DNA. Because of the redundancy of the genetic code, i.e., more than one coding nucleotide triplet (codon) can be used for most of the amino acids used to make proteins, different nucleotide sequences can code for a particular amino acid. Thus, the genetic code can be depicted as follows:

Amino Acid	Codon	Amino Acid	Codon
Phenylalanine (Phe)	TTK	Histidine (His)	CAK
Leucine (Leu)	XTY	Glutamine (Gln)	CAJ
Isoleucine (Ile)	ATM	Asparagine (Asn)	AAK
Methionine (Met)	ATG	Lysine (Lys)	AAJ
Valine (Val)	GTL	Aspartic acid (Asp)	GAK
Serine (Ser)	QRS	Glutamic acid (Glu)	GAJ
Proline (Pro)	CCL	Cysteine (Cys)	TGK
Threonine (Thy)	ACL	Tryptophan (Trp)	TGG
Alanine (Ala)	GCL	Arginine (Arg)	WGZ
Tyrosine (Tyr)	TAK	Glycine (Gly)	GGI
Termination signal	TAJ		

Key: Each 3-letter deoxynucleotide triplet codon corresponds to a trinucleotide of mRNA, having a 5'-end on the left and a 3'-end on the right. All DNA sequences given herein are those of the strand whose sequence corresponds to the mRNA sequence, with thymine substituted for uracil. The letters stand for the purine or pyrimidine bases forming the deoxynucleotide sequence as follows:

A = adenine; G = guanine; C = cytosine; T = thymine

X = T or C if Y is A or G

X = C if Y is C or T

Y = A, G, C or T if X is C

Y = A or G if X is T

W = C or A if Z is A or G

W = C if Z is C or T

Z = A, G, C or T if W is C

Z = A or G if W is A

QR = TC if S is A, G, C or T;

QR = AG if S is T or C

J = A or G

K = T or C

L = A, T, C or G

M = A, C or T

The above shows that the amino acid sequence of the instant plant pathogenesis-related proteins can be prepared using different nucleotide sequences encoding the same amino acid sequence of the proteins. Accordingly, the scope of the present invention includes such "equivalent nucleotide sequences."

cDNA sequences that hybridize with a given enumerated sequence and encode a polypeptide or protein having at least some degree of activity of the corresponding plant pathogenesis-related protein are those which exhibit substantial sequence homology, as defined hereinabove, with the enumerated Sequence such that it hybridizes with the latter under low stringency conditions. Such conditions are described in Examples 17 and 20 of European patent application EP-A-0 392 225. Proteins translated from these hybridizable cDNA sequences have different primary structures from proteins translated from the enumerated Sequences. However, their respective secondary structures are the same. Thus, the former proteins retain some degree of the biological activity (plant disease resistance) of the latter proteins and are considered as plant pathogenesis-related proteins also, e.g. "a PR-O protein," "a PR-N protein," etc. As used elsewhere herein, the term "activity" is referred to as the property of enhanced disease resistance or tolerance conferred upon a transgenic plant by virtue of its ability to express the DNA encoding the PR protein

It is contemplated by the invention that increased disease resistance or tolerance or increased scope of protection against a broader range of pathogens can be achieved by providing for the expression of more than one anti-pathogenic DNA sequence in a transgenic plant or plant tissue. The increased disease resistance thus achieved is due to a synergistic effect created by expressing a combination of anti-pathogenic sequences in the

same plant or plant tissue. Thus, a method is provided which improves protection of a plant against a pest comprising transgenically expressing in said plant two or more DNA molecules encoding anti-pathogenic proteins, wherein the transgenically expressed proteins exert a synergistic effect. Preferably the DNA molecules encode anti-pathogenic proteins according to the invention. Due to synergy, the anti-pathogenic effect of a particular sequence may be more readily detected when combined with one or more other anti-pathogenic sequences than when present alone.

The present invention further includes chimeric genes and vectors comprising one or more anti-pathogenic sequences analogous to the chimeric genes and vectors comprising chemically regulatable sequences described in parts F and G above. Any desired promoter may be used in association with the anti-pathogenic sequences of the invention, including promoters which confer constitutive expression, promoters which confer expression in a particular tissue or subset of tissues and/or at a particular developmental stage, and promoters which confer chemically regulatable expression containing the noncoding sequences taught herein.

Also included as part of the present invention are plant tissues, plants and seeds comprising anti-pathogenic sequences analogous to those plant tissues, plants and seeds comprising chemically regulatable sequences described in part H above.

#### DIFFERENTIAL CLONING AND SCREENING TECHNOLOGY

A method has been conceived and developed which will allow efficient enrichment of sequences present in one population of molecules in greater amounts than in another population. The method's greatest utility is in situations where the populations are very similar and the differentially present sequences represent a very small proportion of the population.

If two populations of clones are similar and one wishes to isolate those clones which are present in one population in higher amounts (i.e. "induced" or differentially regulated"), past techniques involved screening with probes from the two populations (+/- screening; St. John and Davis, *Cell* 16:443-452 (1979)), or enrichment of probes or mRNAs by hybridization and hydroxyapatite (HAP) chromatography (Davis, *et al.*, *Proc. Natl. Acad. Sci. USA* 81: 2194-2198 (1984)). The first method has a demonstrated sensitivity limitation

in that only clones present in greater than about one in 2,000 will be detected. The second is laborious, technically difficult, and achieves enrichments of 20-50 fold at best.

The present method involves exploiting two recent developments in molecular technology: the polymerase chain reaction (Saiki *et al.*, *Science* 239:487-491 (1988)) and biotin-avidin chromatography [Stahl, *et al.*, *Nuc. Acids. Res.* 16: 3026-3038 (1988)]. The polymerase chain reaction (PCR) allows simple synthesis of large amounts of DNA of specified sequence. Biotin-avidin chromatography allows the efficient separation of molecules bearing a biotin affinity tag from those molecules which do not bear the tag.

In its general form, the technique consists of isolating single strands of cDNA representing two different populations ("induced" vs "uninduced"), but of opposite cDNA polarity for the two populations, i.e. one of "sense" polarity relative to mRNA's, and the other its complement, or "anti-sense", polarity relative to mRNA's. The isolated strands from the "induced" population would have no affinity tag, while the strands of opposite polarity from the "uninduced" populations would have stable affinity tags. When these two populations are hybridized together, hybrids will form between complementary strands present in the two populations. Those strands from the "induced" population which have no counterparts, or many fewer counterparts, in the "uninduced" population, remain single stranded.

Due to the presence of the affinity tag (in essence a handle) on the strands of the "uninduced" population molecules, those strands and, most importantly, any hybrid molecules can be removed from the mixture by affinity chromatography. This leaves only those "induced" molecules which are not significantly represented in the "uninduced" population. These "induced" molecules can then be cloned by standard means and serve as an enriched population from which to isolate "induced" clones; alternatively, the enriched molecules can be amplified individually and sequenced directly.

An alternate scheme is the same as described above except that it involves incorporating a labile affinity tag only on the "induced" population molecules, while the affinity tag on the "uninduced" population is stable. "Labile" in this case means that the affinity tag can be removed at will, or be altered at will in such a way that it no longer serves as an affinity tag. In this scheme all the molecules in the hybridization mixture could bind to the affinity matrix, but only those "induced" molecules that are not hybridized to a complementary "uninduced" counterpart could be selectively recovered from the matrix for subsequent cloning.

The advantage of the methods of the invention described above over those previously described is the ability to isolate those genes which are turned on only to low levels, in specific circumstances, and which may play a causative role in some important biological phenomenon.

The present invention teaches the cloning of SAR genes by differential screening of tissues induced and non-induced to the systemic acquired response SAR induction causes the transcription of genes in a protein synthesis-dependent fashion and also a protein synthesis-independent fashion. Two methods were used to clone specifically genes whose induced transcription is protein-synthesis independent. Firstly, cDNAs which were cloned by standard differential screening techniques were further screened on SAR-induced RNA isolated with and without cycloheximide (CHX) pre-treatment. Secondly, a PCR-based "differential display" technique was used to identify SAR-induced, but protein synthesis independent cDNAs directly. Differential display RNAs were prepared with and without SAR induction and CHX treatment. The use of CHX as an inhibitor of protein synthesis is well known in the art and is described by Greenberg *et al.*, *Mol. Cell. Biol.* 6: 1050-1057 (1986), Lau and Nathans, *Proc. Natl. Acad. Sci.* 84: 1182-1186 (1987), and Uknnes *et al.*, *Plant Cell* 5: 159-169 (1993). Thus, a number of genes were cloned which were induced by the SAR response, yet expressed independently of protein synthesis. These cloned genes are likely signal transducers in the pathway leading from induction to the development of the resistant state.

**EXAMPLES*****Example (i): Ligation in Agarose***

Following restriction digestion of plasmid DNA and electrophoretic separation of the fragments on a low melting TAE gel, the bands containing appropriate fragments are precisely excised and heated to 65°C to melt the agarose 2-5 µl are added to 15 µl water and the solution is left at 65°C for 10 minutes. This solution is cooled to 37°C and left for five minutes to equilibrate to temperature 2 µl of 10 X ligase buffer (200 mM Tris, pH 8.0, 100 mM MgCl<sub>2</sub>, 100 mM DTT, 10 mM ATP) are added along with 1 µl T4 DNA ligase (New England BioLabs), and this solution is allowed to solidify and incubate at 15°C overnight.

***Example (ii): Transformation From Agarose***

The agarose containing the appropriate DNA is melted by incubating at 65°C for 10 minutes 10 µl of this solution are added to 30 µl of TE buffer (10 mM Tris pH 7.5, 1 mM EDTA), mixed and allowed to stand at room temperature. Frozen competent cells (*E. coli* strain DH5) are placed on wet ice to thaw. The diluted DNA solution is added to 200 µl of cells and allowed to stand on ice for 20 minutes. The cells containing the DNA are then heat-shocked for 90 seconds at 41°C. The cells are then left at room temperature for 10 minutes 0.8 ml of SOC medium (Hanahan, D., *J. Mol. Biol.* 166: 557-580 (1983)) is added and the culture is incubated at 37°C for one hour 100 µl of the culture is plated on LB plates (Miller, *supra*) containing 100 µg/ml ampicillin (L-amp) and the plates are incubated overnight at 37°C. Positive colonies are picked and restreaked to a second L-amp plate and the plates are incubated overnight at 37°C.

***Example (iii): Labelling DNA Restriction Fragments***

DNA is treated with the appropriate restriction enzymes and fragments are separated by electrophoresis on a low-gelling temperature agarose gel. A band containing the fragment of interest is excised and the DNA purified by standard techniques 50 ng of the DNA fragment is labelled using the IBI Random primer kit "Prime time" according to the manufacturers directions.

- 25 -

**Example (iv): Southern Blotting**

3 µg of tobacco DNA is digested with various restriction enzymes under the conditions suggested by the supplier. The DNA is extracted with phenol, precipitated with ethanol and then resuspended in gel loading buffer (15% ficoll, 0.025% bromophenol blue, 10 mM EDTA, pH 8). Samples are loaded and electrophoresed on a 0.5% agarose gel at 5 V/cm until the bromophenol blue dye reaches the end of the gel. The DNA is transferred to Gene-Screen Plus (DuPont) using the alkaline transfer procedure as described by the supplier. Pre-hybridization, hybridization and washing are according to the manufacturer's recommendation. Hybridization is detected by autoradiography.

**Example (v): Molecular Adapters**

A typical molecular adaptor for the conversion of a PstI site to a BamHI site is the sequence

5'-GGGATCCCTGCA-3' (SEQ ID No. 22).

This molecule is synthesized on an Applied Biosystems Synthesizer using B-cyanoethylphosphoramidite chemistry and purified by reverse-phase HPLC. About 2 µg of this oligonucleotide is kinased according to Maniatis *et al.*, *supra*, p. 125. The oligonucleotide solution is heated to 65°C in a water bath and allowed to cool to room temperature over a period of about 30 minutes. An approximately 10-fold molar excess of this annealed adapter is added to the digested DNA along with 10 X ligase buffer, T4 DNA ligase, and an appropriate amount of water. A typical reaction is:

DNA to be adapted:	1-2 µl (~ 1 pmol)
Adapter:	1 µl (~10 pmol)
10 X ligase buffer:	1 µl
T4 DNA ligase:	1 µl
Water:	5-6 µl

This solution is incubated at 12-15°C for 30 minutes, and heated to 65°C for 30 minutes to inactivate the ligase. The salt concentration and volume are adjusted for the appropriate restriction digest and the adapted DNA is digested to expose the adapted

"sticky end." Unincorporated adapters are removed either by electrophoresis on an agarose gel or by sequential isopropanol precipitations.

**Example (vi):      *Primer Extension Mapping***

**A. Synthesis and 5' End Labeling of Primers for Primer Extension**

The following primer oligomers are synthesized using an Applied Biosystems Synthesizer and  $\beta$ -cyanoethylphosphoramidite chemistry:

PR-1: 5'-ATAGTCTTGTGAGAGTT-3' (SEQ ID No. 23)

GUS: 5'-TCACGGGTTGGGGTTCTAC-3' (SEQ ID No. 24)

AHAS: 5'-AGGAGATGGTTGGTGGA-3' (SEQ ID No. 25)

BT: 5'-ATACGTTCTACTATCATAGT-3' (SEQ ID No. 26)

The oligonucleotides are purified by reverse-phase high pressure liquid chromatography (HPLC) 5 pmol of each oligo is kinased (Maniatis, T. *et al.*, *supra*, at p. 125) using 200  $\mu$ C of  $^{32}$ P-ATP (6000 Ci/mmol, 10  $\mu$ Ci/ $\mu$ l). After incubation at 37°C for 30 minutes, the reaction is diluted to 100  $\mu$ l, extracted with phenol/chloroform and then precipitated three times with 50  $\mu$ g carrier RNA. The final precipitate is resuspended in 1 X reverse-transcriptase buffer (50 mM Tris-HCl, pH 7.5, 40 mM KCl, 3 mM MgCl<sub>2</sub>) at a concentration of 2 nM. The specific activity of the labeled oligonucleotide is determined to be about 3 X 10<sup>6</sup> Ccpm/pmol.

**B. Total RNA Preparation**

Total RNA is prepared essentially as described by Lagrimini, L.M. *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 7542 (1987). Tissue is ground under liquid nitrogen in a mortar and pestle. The ground tissue is added to grinding buffer (Lagrimini *et al.*, *supra*) using 2.5 ml per gram tissue. An equal volume of phenol is then added and the emulsion is homogenized in a Brinkman polytron. A one-half volume of chloroform is added and the emulsion is gently mixed for 15 minutes. The phases are separated by centrifugation and

the aqueous phase is removed. RNA is precipitated by the addition of sodium acetate to 0.3 M and 2.5 volumes ethanol. The precipitate is collected by centrifugation and resuspended in 2 ml sterile water. Lithium chloride is added to a final concentration of 3 M and left at 4°C overnight. The precipitate is collected by centrifugation and the pellet is washed with ice-cold 80% ethanol. The pellet is dried and resuspended in 500 µl sterile water. The concentration of this total RNA preparation is determined spectrophotometrically.

Alternatively, RNA is extracted from callus as described above except that the callus tissue is cut into cubes approximately 3 mm in size, and added to pre-chilled mortars and pestles for grinding in liquid nitrogen prior to the polytron step.

#### C. Primer Extension

50 µg of total RNA is lyophilized in a 500 µl Eppendorf tube. The RNA is resuspended in 30 µl of radiolabeled probe solution and heated to 70°C for 10 minutes. The tube is slowly cooled to 37°C and allowed to incubate overnight. Without removing the tube from the 37°C water bath, 2 µl of 10 X reverse-transcriptase buffer (500 mM Tris-HCl, pH 7.5, 400 mM KCl, 30 mM MgCl<sub>2</sub>), 1 µl 5 mg/ml bovine serum albumin, 5 µl 100 mM dithiothreitol, 5 µl 10 X dNTPs (10 mM of each dNTP in H<sub>2</sub>O), 3 µl H<sub>2</sub>O, 2 µl RNasin (80 units), and 2 µl reverse transcriptase (400 units) are added and the reaction is incubated at 37°C for 30 minutes. To stop the reaction, 5 µl of 3 M sodium acetate, pH 5, and 150 µl absolute ethanol are added. The tube is left at -20°C for 30 minutes, the precipitate is collected by centrifugation, washed with 80% ethanol and allowed to air-dry. The precipitate is resuspended in 10-20 µl of loading dye (90% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol, 1 mM EDTA) and the extension products are separated on a 6% sequencing gel (Maniatis, T. *et al.*, supra). Extension products are visualized by autoradiography.

#### ***Example (vii): S1 Nuclease Mapping***

The plasmid pBS-PR1013Cla is digested with SfaNI, dephosphorylated with calf intestinal phosphatase and kinased with <sup>32</sup>P-ATP. Following phenol extraction and ethanol precipitation, the DNA is digested with BstEII and the 300 bp fragment from 750 to 1035 of Figure 1 is isolated after electrophoresis on a low gelling temperature agarose gel. The probe is resuspended in formamide hybridization buffer (Berk, A.J. *et al.*, Cell 12, 721

(1977)) at a concentration of about 2 nM. The specific activity of the probe is about  $5 \times 10^{10}$  Ccpm/pmol.

Lyophilized, total RNA (50 µg) is dissolved in 30 µl of the probe solution, and the tubes are heated to 65°C for 10 minutes, then allowed to hybridize overnight at 48°C. S1 nuclease treatment and gel electrophoresis are essentially as described, using an S1 concentration of 400 units/ml and an incubation temperature of 30°C. The appropriate S1 nuclease concentration is determined in pilot experiments.

**Example (viii): *Mapping the Transcriptional Start Site***

The transcriptional start site for the PR-1a gene is determined by a combination of S1 nuclease mapping and primer extension analysis. An autoradiogram of a primer extension experiment using either RNA isolated from TMV-infected leaves or an mp19 subclone of the Xhol-PstI fragment as a template and a 17 base oligonucleotide complementary to positions 1025 to 1042 of the PR-1a sequence as a specific primer is examined. The primer itself is labeled at the 5' phosphate, therefore the size of the extension product will be identical to the size of the corresponding band in the sequencing lanes. The appearance of two strong bands corresponding to positions 902 and 903 and a weak band at position 901 of the genomic clone suggests transcription initiating at either of these positions. However, primer extension analysis alone cannot be used to identify the 5' end of a mRNA. For instance, the mRNA may contain a 5' end that has been spliced from an upstream location.

To determine conclusively the 5' end, high resolution S1 nuclease mapping is used in conjunction with primer extension. An SfaNI fragment is labeled at the 5' end and digested with BstEII to yield a strand specific probe extending from position 759 to 1040. This probe is used to map the 5' end of PR-1a transcripts in RNA isolated from TMV-infected tobacco leaves. A major band of 137 2 bases is found which corresponds to positions 901 to 905 of the genomic clone. In high resolution S1 experiments, where the digestion products are electrophoresed along with a size standard of sequencing reactions performed on the probe, three bands are visualized corresponding to positions 901, 902 and 903. These results confirm the primer extension analysis and place the 5' end of the PR-1 mRNA at either position 901, 902 or 903. With regard to transcription initiation, one possible interpretation of these results is that RNA polymerase begins transcription at either base 901, 902 or 903 with more or less equal probability. However, since eukaryotic

transcription favors initiation at an A, a more likely explanation for the apparent multiple 5' ends is that the PR-1a mRNA begins at position 903 (an A) and the PR-1b and -1c mRNAs begin each at one of the other positions on their corresponding genes.

***Example (ix): General Techniques for Peptide Generation, Purification, and Automated Sequencing***

The proteins relevant to these examples are isolated, purified and sequenced in some cases, for the first time and in accordance with literature procedures in other, for the purpose of allowing the isolation of the corresponding cDNA's and ultimately for confirming the identities of their cDNA's and chemically inducible genes.

**A. Reduction and Alkylation**

Purified, lyophilized protein is dissolved in 6 M guanidine-HCl containing 1 M Tris-HCl, pH 8.6, 10 mM EDTA. Dithiothreitol is added to 20 mM and 4-vinylpyridine is added to a final concentration of 50 mM. The sample is then incubated for 1.5 hours under nitrogen. The pyridylethylated material is desalted on HPLC using an Aquapore phenyl column (2.1 x 10 cm, Brownlee). The column is eluted with a linear, 5-80% gradient of acetonitrile/isopropanol (1:1) in 0.1% trifluoroacetic acid (TFA).

**B. Cyanogen Bromide Cleavage and Removal of Pyroglutamate**

Cyanogen bromide cleavage is performed *in situ* according to Simpson, R.J. et al., *Biochem. Intl.* 8: 787 (1984) Digestion of PR-1 protein with pyroglutamate aminopeptidase (Boehringer Mannheim) is carried out according to Allen, G., *Plant Sci. Lett.* 26: 173 (1982).

**C. LysC digestion**

Protein is digested with endoproteinase Lys-C (Boehringer Mannheim) in 0.1 M Tris-HCl, pH 8.5, for 24 hours at room temperature using an enzyme:substrate ratio of 1:10. Resulting peptides are isolated by HPLC using an Aquapore C-8 column (1 x 22 cm, Brownlee) eluted with a linear acetonitrile/isopropanol (1:1 ratio) gradient (0 to 60%) in 0.1% TFA.

#### D. Trypsin Digestion

Digestion with trypsin (Cooper) is performed in 0.1 M ammonium bicarbonate, pH 8.2, containing 0.1 M calcium chloride for five hours at 37°C using an enzyme:substrate ratio of 1:100. Peptides generated are separated on HPLC using the same conditions as with the Lys-C peptides or performed in 0.1 M Tris-HCl pH 8.5 for 24 hours at 37°C using an enzyme to substrate ratio of 1:50. Peptides are isolated by HPLC using a Vydac C-18 column (2.1 x 150 mm) with a linear 0 to 60% acetonitrile:isopropanol (1:1) gradient in 0.1 % TFA.

#### E. Sequencing

Automated Edman degradations are performed with an Applied Biosystems 470A gas-phase sequencer. Phenylthiohydantoin (PTH) amino acids are identified using an Applied Biosystems 120A PTH analyzer.

#### *Example 1A: Cloning of cDNAs corresponding to SAR CHX-independent genes from tobacco*

A number of cDNAs were cloned by differential screening from cDNA prepared from induced and non-induced tissue. The induced cDNA was prepared from tobacco leaves which had been pre-treated with methyl benzo-1,2,3-thiadiazole-7-carboxylate (BTH), whereas the non-induced cDNA was prepared from tissue which had not been pre-treated with BTH. cDNA libraries were prepared in λZAP II (STRATAGENE). A standard differential screening technique was used. Plaques carrying induced cDNA were plated at low density and transferred to two sets of hybridization filters. Known SAR gene sequences were hybridized to the first filter and uninduced cDNA to the second. The second filter was then stripped and hybridised with induced cDNA. Plaques which hybridized with the induced cDNA probe, but not with the uninduced cDNA probe or with known SAR gene sequences were potential novel SAR genes and were picked directly for further analysis. Plating had been at a plaque density which was sufficiently low to enable these plaques to be picked as nearly pure plaques. Multiple candidates for each plaque were *in vivo* excised, according to the manufacturer's recommended conditions, for further screening by Northern hybridization to RNA isolated from either untreated or BTH-treated tobacco plants. Individual clones chosen from the secondary screen were further analyzed by Northern

hybridisation to RNA isolated from tobacco plants which had been pre-treated with salicylic acid (3 mM), INA (1 mM), and BTH (1 mg/ml), all in the presence or absence of cycloheximide (CHX; 1 mg/ml). Inducer pre-treatments were done at 2 h, one day, and eight days before the isolation of RNA, whereas CHX treatment was done at one day before isolation of RNA. Three cDNAs were found induced in a protein synthesis independent fashion. The genes corresponding to these cDNAs have been designated p66B1 (SEQ ID NO. 14), p11.30.13 (SEQ ID NO. 11), and p1.4.3 (SEQ ID NO. 12 / SEQ ID NO. 13) and are likely signal transducers of the SAR response. Gene p1.4.3 has previously been disclosed as a thioredoxin (Brugidou *et al.*, *Mol. Gen. Genet.* 238: 285-293 (1993)). However, this is the first disclosure of the gene's likely involvement in the systemic acquired response. Three cDNAs were found to be expressed in a protein synthesis dependent fashion. These were designated p1.1.1, p11.31.4, and p14.22.3 and are listed as SEQ. ID Nos. 9, 10, and 15, respectively.

***Example 1B: Cloning of cDNAs corresponding to SAR CHX-independent genes from *Arabidopsis* using INA for induction***

Total RNA was isolated from the following *Arabidopsis* lines: (1) untreated, (2) INA treated (0.25 mg/ml), (3) CHX treated (1 mg/ml), and (4) INA + CHX treated. Treatments were made 1 day before RNA isolation. RNA thus isolated was subjected to "differential display" using the protocol described by Liang and Pardee, *Science* 257: 967-971 (1992). Amplified fragments which were found in both the INA as well as the INA + CHX treated RNA samples were gel-purified and used as probes on Northern blots carrying similarly induced RNA samples. Fragments for which Northern hybridization confirmed the induction profile apparent from differential display were subcloned into a plasmid vector. Using the cloned fragment it was possible to isolate near full-length cDNAs from a cDNA library produced by BTH induction:

Mature *Arabidopsis thaliana* ecotype Columbia (Lehle Seeds, Tucson, AZ) plants are sprayed with a 0.5 mg/ml suspension of a wettable powder formulation of methylbenzo-1,2,3-thiadiazole-7-carboxylate consisting of 25% active ingredient. Seven days later, the leaf tissue is harvested and frozen in liquid N<sub>2</sub>. Total RNA is isolated as described in Example (vi) and Poly (A)<sup>+</sup> RNA is isolated using a Poly (A) Quik<sup>TM</sup> mRNA isolation kit from Stratagene (La Jolla, CA). This Poly (A)<sup>+</sup> RNA is then used to make a cDNA library in the

uni-zap<sup>TM</sup> XR vector (Stratagene) using a ZAP-cDNA<sup>TM</sup> Gigapack<sup>R</sup> II Gold cloning kit from Stratagene. A portion of the cDNA library is amplified as described in the Stratagene kit.

The cDNA pDPA2 was cloned using this technique and is induced by the SAR response in a protein synthesis independent fashion. Its sequence is listed in Seq. ID No. 16.

***Example 1C: Cloning of cDNAs corresponding to SAR CHX-independent genes from *Arabidopsis* using BTH for induction***

RNA isolated from (1) untreated, (2) BTH treated, and (3) BTH + CHX treated *Arabidopsis* lines was subjected to "differential display" using the protocol described by Liang and Pardee, *Science* 257: 967-971 (1992) and as described in example 1B above. Amplified fragments which were found in both the BTH as well as the BTH + CHX treated RNA samples were gel-purified and used as probes on Northern blots carrying similarly induced RNA samples. Fragments for which Northern hybridization confirmed the induction profile apparent from differential display were subcloned into a plasmid vector. The following clones induced in a protein synthesis-independent manner were obtained:

- PSI-1 (SEQ ID NO. 33): cDNA homologous to EMBL Accession #Z26429, an *A. thaliana* transcribed sequence (clone FAFC80-1; 5' end and actin depolymerizing factor from *L. longiflorum*), and EMBL Accession #Z14110, an *L. longiflorum* mRNA for actin depolymerizing factor
- PSI-2 (SEQ ID NO. 34): cDNA with no known homology
- PSI-3 (SEQ ID NO. 35): cDNA homologous to EMBL Accession # T20772, *A. thaliana* cDNA 92H8T7, and EMBL Accession #T04101, *A. thaliana* cDNA clone SCG7T7P
- PSI-4 (SEQ ID NO. 36): cDNA homologous to EMBL Accession #D22118, a partial Rice cDNA sequence
- PSI-5 (SEQ ID NO. 37): cDNA with no known homology

***Example 2A: Cloning of Class IV Chitinase cDNAs***

To clone recombinant or chimaeric DNA molecules encoding class IV chitinase class specific degenerate oligonucleotides were designed from areas of homology between bean PR4 chitinase (Margis-Pinheiro *et al.*, *Plant Mol. Biol.* 17: 243-253 (1991)) and sugar beet chitinase 4 (Mikkelsen *et al.*, in "Advances in Chitin and Chitosan, ed. by Brine *et al.*, pub. by Elsevier, Amsterdam (1992)): oligonucleotides were designed degenerate for the peptide

sequences HFCYIEE (forward, spanning nucleotides 406-426; SEQ ID NO. 27), and IRAING (reverse, spanning nucleotides 705 to 675; SEQ ID NO. 28). DNA was extracted from two-week old plants of *Arabidopsis thaliana* ecotype Landsberg and amplified using a Perkin-Elmer thermal cycler 480 at the following cycle settings: 94°C for 5 minutes; 35 cycles at 94°C, 1 minute, 43°C or 45°C, 1 minute, and 72°C, 2 minutes; followed by 5 minutes at 72°C. The amplified fragment was gel purified, collected by centrifugation through Whatman paper, ethanol precipitated, resuspended in TE, digested with BamHI and NsI and subcloned into pTZ18U (Pharmacia). Four clones of the fragment were sequenced; they differed only within the oligo-derived sequence as could be expected from amplification with degenerate oligonucleotides.

Two different cDNAs were isolated simultaneously by screening a leaf tissue cDNA library (Uknes *et al.*, *Plant Cell* 4: 645-656 (1992)) at high stringency with the PCR amplified genomic fragment described above. Duplicate plaque lifts were taken with nitrocellulose filters (Schleicher & Schuell, Keene, NH) (Ausubel *et al.*, in "Current Protocols in Molecular Biology, pub. by J. Wiley & Sons, New York (1987)"). Probes were labelled by random priming (using the labeling system supplied by Gibco BRL, Gaithersburg, MD). Hybridization and washing were done at 65°C according to Church and Gilbert, *Proc. Natl. Acad. Sci.* 81: 1991-1995 (1984). Positive plaques were purified and plasmids containing the cDNA inserts were *in vivo* excised for DNA sequence determination. Of the six positive clones two contained an insert with structural homology to previously characterized class IV chitinases and were designated class IV chitinase type A and four clones were divergent in that they lacked the class IV chitinase hevein domain; these were designated class IV chitinase type B. As none of the cDNA inserts was full-length, an additional 29 bp of class IV chitinase type A and 17 bp of class IV chitinase type B, both containing a methionine initiation codon were amplified from ethephon-induced RNA using the 5' RACE system for rapid amplification of cDNA ends (Gibco BRL, Gaithersburg, MD). Sequence comparisons were performed using the GAP and PILEUP features of the Genetics Computer Group software (Genetics Computer Group, Madison, WI).

Class IV chitinase type A is a 1079 base pair cDNA with an open reading frame of 264 amino acids containing the characteristic cysteine-rich hevein and chitinolytic domains and the three short deletions typically found in class IV chitinases (see SEQ ID No. 1). The cDNA for class IV chitinase type B is 952 base pairs in length and encodes a protein of 214 amino acids which lacks a hevein domain and contains a fourth deletion (see SEQ ID No.

2). The cDNAs are 71% identical overall and 80% identical over coding sequence. The predicted protein encoded by *Arabidopsis* class IV chitinase type A is 89% homologous to the basic *Brassica napus* (rapeseed) class IV chitinase, 61% homologous to basic *Beta vulgaris* (sugar beet) class IV chitinase, 57% homologous to basic *Zea mays* class IV chitinase B, 58% homologous to acidic *Phaseolus vulgaris* (bean) PR4 class IV chitinase, and 55% homologous to acidic *Dioscorea japonica* (yam) class IV chitinase. It is 42% homologous to *Arabidopsis thaliana* basic class I chitinase.

The predicted mature protein encoded by *Arabidopsis* class IV chitinase type A has a molecular weight of 25695 D and a pI of 7.8; whereas the protein encoded by *Arabidopsis* class IV chitinase type B has a molecular weight of 20553 D and a pI of 10 assuming the removal of a signal peptide based on homology to tobacco PR-Q.

Northern analysis showed that both chitinases were induced by TCV infection, confirming their classification as PR-proteins.

Using techniques well known in the art, these cDNAs can be cloned into expression cassettes and vectors for transfer to transgenic plants. Typical techniques used in the art are described in section 4A (examples 22 to 35), section 5A (examples 36 to 56) and section 6 (examples 57 to 80) of EP-A 0 392 225.

***Example 2B: Isolation of Chemically Induced Wheat cDNAs Using the Method of Differential Plaque Filter Hybridization***

Recombinant or chimaeric DNA molecules comprising the cDNA of a wild-type wheat gene which wild-type gene can be chemically induced in a plant were obtained in the following way:

Samples of winter wheat (cultivar Kanzler) were harvested 2-3 days after treatment with either water or 200 ppm of the plant activator compound benzo-1,2,3-thiodiazole-carboxylic acid. Total RNA was prepared from frozen tissue samples using a standard phenol extraction/LiCl precipitation procedure (Lagrimini et al., *Proc. Natl. Acad. Sci.* 84: 7542-7546 (1987)). PolyA (+) RNA was purified from total RNA using the Poly(A) Quik mRNA purification kit (Stratagene Cloning Systems, LaJolla, CA). A bacteriophage lambda ZAP II cDNA library was prepared from the benzo-1,2,3-thio-diazolecarboxylic acid treated polyA(+) sample using the Uni-Zap XR Gigapack II Gold cloning kit (Stratagene) as described by the manufacturer. The phage library was plated at a density of approximately 5000 plaques on a 10 cm petri dish and grown for 6-8 hours at 37°C. Duplicate filter lifts of

the plaques were made using nitrocellulose membranes (Schleicher & Schuell, Keene, NH). Labelled first strand cDNA probes were prepared from polyA of both the water control and the benzo-1,2,3-thiodiazolecarboxylic acid-treated samples using  $^{32}\text{P}$ -dCTP and the AMV reverse transcriptase (GibcoBRL, Gaithersburg, MD) under the manufacturer's conditions. Each probe was hybridized ( $>106$  cpm/ml) with one set of the duplicate lifts overnight at 65°C. Hybridization and wash conditions were as described in Church and Gilbert, *Proc. Natl. Acad. Sci. 81: 1991-1995* (1984). Hybridization was detected by autoradiography.

Plaques appearing to hybridize preferentially to the chemically treated cDNA were purified and their cDNA inserts were amplified using the GeneAmp Polymerase Chain Reaction (PCR) kit (Perkin Elmer, Norwalk, CT) and primers homologous to the flanking lambda Zap II sequences. The amplified inserts were excised from a low melting temperature SeaPlaque GTG agarose gel (FMC BioProducts, Rockland, ME) and labelled using  $^{32}\text{P}$ -dCTP and the Random Primers DNA Labeling System (Gibco BRL). These probes were hybridized with total RNA blots (Ausubel *et al.*, in "Current Protocols in Molecular Biology", pub. by J. Wiley & Sons, New York (1987)) of control and benzo-1,2,3-thiodiazolecarboxylic acid treated RNAs to verify that they contained chemically induced cDNAs.

The induced clones were *in vivo* excised into pBluescript plasmids according to the manufacturer's instructions (Stratagene) and plasmid DNAs were purified using Magic Miniprep columns (Promega Biotech, Madison, WI). The cDNA sequences were determined by the chain termination method using dideoxy terminators labelled with fluorescent dyes (Applied Biosystems, Inc., Foster City, CA). The DNA and the predicted amino acid sequences were compared to available databases using the GAP (Devereaux *et al.*, *Nucl. Acids Res. 12: 387-395* (1984)) and the BLAST (Altschul *et al.*, *J. Mol. Biol. 215: 403-410* (1990)) programs.

The nucleotide sequence of one apparently full length induced clone is set forth in SEQ ID No. 3. This clone, denoted WCI-1 (Wheat Chemical Induction), was found to share limited homology with two rice cDNAs, one expressed specifically in the shoot apical meristem (De Pater and Schilperoort, *Plant Mol. Biol. 18: 161-164* (1992)), the other inducible by salt stress (Claes *et al.*, *Plant Cell 2: 19-27* (1990)). The function of these proteins is unknown.

The predicted amino acid sequence of a second highly induced cDNA, WCI-2, clearly identified it as an isozyme of wheat lipoxygenase, based on its homology to other

plant lipoxygenases in the database. The DNA sequence of this cDNA is shown in SEQ ID No. 4.

A third class of induced cDNA (WCI-3) was isolated which to date shows no significant homology to sequences in the databases. The DNA sequence of this apparently full-length clone is set forth in SEQ ID No. 6.

Using techniques well known in the art, these cDNAs can be cloned into expression cassettes and vectors for transfer to transgenic plants. Typical techniques used in the art are described in section 4A (examples 22 to 35), section 5A (examples 36 to 56) and section 6 (examples 57 to 80) of EP-A 0 392 225.

***Example 2C: Isolation of Wheat cDNAs Specifically Induced by Treatment With benzo-1,2,3-thiodiazolecarboxylic acid, Using the Method of Differential cDNA Display***

The total and polyA(+) samples described in Example (vi) were used for PCR differential display of the mRNA from water and benzo-1,2,3-thiodiazolecarboxylic acid treated wheat tissue essentially as described in Liang and Pardee, *Science* 257: 967-971 (1992). Amplified cDNA fragments that appeared to be present only in the chemically treated sample were excised from the dried sequencing gel and electroeluted using a Centriliutor device and Centricon-30 Microconcentrators (Amicon, Beverly, MA). The purified fragments were PCR amplified using primers that consisted of the original differential display 10-mers plus an additional 10 bases of unique sequence added to their 5' ends. After 8 PCR cycles at low annealing temperature (42-45°C), the annealing temperature was raised to 60°C for an additional 30 cycles. Essentially 100% of the fragments could then be visualized by EtBr staining and were excised from a SeaPlaque GTG agarose gel (FMC).

These gel fragments were labelled and used to probe RNA blots as previously described. Fragments that hybridized only with chemically treated RNA were TA-cloned into the plasmid vector pCR II using a TA Cloning Kit (Invitrogen Corporation, San Diego, CA). Inserts from the plasmids were screened again against control/chemical RNA blots to verify that the inducible gene fragment had been subcloned. Induced fragments were then labelled by random priming and hybridized against filter lifts of the chemically induced cDNA library as described in Example 2A. Hybridizing plaques were purified, sequenced, and analyzed as previously described in order to obtain full length clones corresponding to the

original small (200-400 bp) fragments, and to identify the induced gene product where possible.

The nucleotide sequence of one clone obtained by this procedure is set forth in SEQ ID No. 7. This clone, WCI-4, has some homology to known thiol protease sequences from a variety of sources and may therefore be a thiol protease.

Partial sequences of an additional induced gene that was isolated using differential display is set forth in SEQ ID No. 8. This fragment shows no database homology and has been designated WCI-5.

Using techniques well known in the art, these cDNAs can be cloned into expression cassettes and vectors for transfer to transgenic plants. Typical techniques used in the art are described in section 4A (examples 22 to 35), section 5A (examples 36 to 56) and section 6 (examples 57 to 80) of EP-A 0 392 225.

***Example 3A: Preparation of cDNA Library From TMV-Infected Tobacco Leaves.***

*Nicotiana tabacum* cv. Xanthi-nc leaves are infected with tobacco mosaic virus and harvested five days post-infection. Total RNA is prepared as described above and poly A+ RNA is isolated using standard techniques. A cDNA library is constructed using this poly A+ RNA in the Lambda *ongC* cloning vector (Stratagene) essentially as described (Gubler, U. and Hoffman, B. J., *Gene* 25: 263 (1983)).

***Example 3B: Isolation of cDNA Clones Encoding Chitinase/Lysozymes***

To clone recombinant or chimaeric DNA molecules encoding chitinase/lysozyme about 300,000 plaques of the TMV-infected tobacco cDNA library described in Example 3A are screened using a labeled cDNA probe encoding the cucumber chitinase/lysozyme cDNA and washing filters at 50°C in 0.125 mM NaCl, 1% SDS, 40 mM sodium phosphate (pH 7.2), 1 mM EDTA. Positive plaques are purified and the DNA sequence of two clones, named pBSCL2 and pBSTCL226 are determined. These are presented in SEQ ID Nos. 20 and 21, respectively. The proteins encoded in the clones of these sequences are determined to be chitinase/lysozymes based on structural homology to the cucumber chitinase/lysozyme.

In addition, a protein is purified from intercellular fluid of TMV-infected tobacco. Peptides are generated and sequenced as described in Example (ix)

The protein encoded by pBSTCL226, corresponding to an acidic isoform of chitinase/lysozyme, was found to match the deduced peptide sequences exactly.

Using techniques well known in the art, these cDNAs can be cloned into expression cassettes and vectors for transfer to transgenic plants. Typical techniques used in the art are described in section 4A (examples 22 to 35), section 5A (examples 36 to 56) and section 6 (examples 57 to 80) of EP-A 0 392 225.

***Example 3C: Analysis of Seed Lines Derived From Transformation of Tobacco with the Basic and Acidic Class III Tobacco Chitinase Genes***

Binary constructions carrying the tobacco basic and acidic class III chitinases were transformed into *Nicotiana tabacum* cv *Xanthi-nc* Leaf tissue samples were taken from T1 plants and assayed for expression of either transgene protein (by Western analysis) or transgene RNA (by Northern analysis) T1 plants found to express the transgene at high levels were advanced to T3 seed lines as described above.

***Example 3D: Analysis of Transgenic Plants Expressing Class III Chitinase for Pest Resistance***

Transgenic plant lines expressing one or more class III chitinase genes are assessed for resistance to numerous pests. Approximately six plants of each line are tested. Pests at the appropriate stage of their growth cycle (such as larvae) are introduced at the appropriate stage of plant development. Plants are later assayed for % of leaf or tissue area eaten, % of introduced larvae surviving, and the weight of surviving larvae.

***Example 3E: Analysis of Transgenic Tobacco Plants Expressing Class III Chitinase for Pest Resistance***

Transgenic tobacco plants expressing class III chitinase are assessed for resistance to numerous pests. Those tested include *Spodoptera exigua* (beet armyworm), the green peach tobacco aphid, *Manduca sexta*, and various nematodes, weevils, mites, and other pests.

**Example 3F: Analysis of Transgenic Plants Expressing Basic Class III Chitinase for Resistance to *Heliothis virescens***

Transgenic tobacco lines expressing the tobacco basic class III chitinase gene were assessed for resistance to the insect *Heliothis virescens*. Eight plants of each line were tested 50 mm leaf discs were cut from the youngest leaves of transgenic plants approximately six weeks after germination and 3 larvae (2nd stage) were allowed to feed on each disc for 3-7 days. After 3-7 days the leaf discs were assessed for area eaten, % of larval survivors and weight (in mg) of larval survivors. Results of three separate experiments are shown in the tables provided below.

The results show elevated resistance to *Heliothis* in basic class III chitinase overexpressing plants when compared to non-transformed control lines *Heliothis virescens* (the tobacco budworm) causes considerable damage in tobacco crops and is particularly recalcitrant to control using pesticides as the larvae burrow deep within the plant. Furthermore, other *Heliothis* species cause similar damage to cotton and other crops.

- 40 -

(1)	%LEAF EATEN	%SURV IVORS	WT (mg) LARVAE	(2)	%LEAF EATEN	%SURV IVORS	WT(mg) LARVAE
-----	----------------	----------------	-------------------	-----	----------------	----------------	------------------

Line

Control	72	100	23		77	100	26
3505 C-15-5	45	92	14		65	96	19
3505 C-5-3	55	96	17		60	92	23
3505 C-4-2	57	92	19		60	100	20
3505 C-12-8	-	-	-		44	87	15

(3)	%LEAF EATEN	%SURV IVORS	WT (mg) LARVAE
-----	----------------	----------------	-------------------

Line

Control	65	100	53
3505 C-12-8	59	100	34

A summary of the two experiments detailed above is given below:

(1) Line:	Control	3505 C-15-5	3505 C-5-3	3505 C-4-2
-----------	---------	-------------	------------	------------

Resistance	0	4	2.5	2
------------	---	---	-----	---

Index

(2) Line:	Control	3505 C-15-5	3505 C-5-3	3505 C-4-2	3505 C-12-8
-----------	---------	-------------	------------	------------	-------------

Resistance	0	2	3	3	4
------------	---	---	---	---	---

Index

(3) Line:	Control	3505 C-12-8
-----------	---------	-------------

Resistance	0	3
------------	---	---

Index

## Resistance Index:

- S = significantly more susceptible than control plants (i.e. wild-type)
- (S) = more susceptible than control plants, but not significantly in all evaluations
- 0 = like wild type plants
- 1 = some signs of resistance but not significantly different to controls
- 2 = identified resistance, significant only in some evaluations
- 3 = resistance with small significant differences in all evaluations
- 4 = high resistance, but symptoms or damage can still be observed
- 5 = total resistance. No symptoms or damage.

***Example 4A: Isolation of a Genomic Clone for Arabidopsis PR-1***

The *Arabidopsis* PR-1 cDNA cloned in pAPR1C-1 (sequence SEQ ID No. 17) was used as a hybridization probe in screening an *Arabidopsis* λEMBL 3 genomic library (purchased from Clontech). Four hybridizing plaques were plaque purified using conventional techniques and λDNA was isolated from each one with λmdasorb (Promega). The λDNA thus isolated was digested with Xhol, electrophoresed, transferred to hybridization membrane for hybridization with the PR-1 cDNA. A fragment of 7 kb was found to hybridize to the cDNA. DNA from one of the four purified plaques was then redigested with Xhol and ligated into the Xhol site of pBluescript (Stratagene). A colony carrying a recombinant or chimaeric DNA molecule comprising the promoter region of the *Arabidopsis* PR-1 gene was identified by probing with oligonucleotide DC21 from the PR-1 coding sequence (position +110 to +84) and the plasmid contained therein was designated pAtPR1-P and deposited January 5, 1994 with the Agricultural Research Culture Collection, International Depositing Authority, 1815 N. University Street, Peoria, Illinois 61604 under the accession no. NRRL B-21169. Restriction analysis identified the 7 kb Xhol fragment as extending 4.2 kb upstream of the ATG of the PR1 gene.

***Example 4B: Fusion of the Arabidopsis PR-1 Promoter to the Firefly Luciferase Gene***

Plasmid pDO432 containing a gene encoding luciferase (LUC) from firefly was received from Dr David Ow (University of California, San Diego; see Ow *et al.*, 1986; Science 234: 856). The LUC gene was excised from pDO432 by digestion with XbaI (at position +45 relative to the ATG) and SstI (approximately 1.8 kb downstream of the ATG and outside the LUC coding region). Additionally, an EcoRI-XbaI promoter fragment was

excised from pAtPR1-P; this fragment was 1.4 kb in size and delineated by an XbaI site 2.8 kb upstream of the PR-1 ATG and an EcoRI site in the pBluescript polylinker distal to the 5' end of the cloned promoter fragment (at -4.2 kb relative to the ATG). These two fragments were cloned by three-way ligation into EcoRI/SacI cleaved pBluescript thus orienting the LUC gene adjacent to the upstream PR-1 promoter fragment (pAtPR1-Q).

Subsequently, pAtPR1-P was cleaved with XbaI (at the -2.8 kb position and within the pBluescript polylinker) and religated to generate a PR-1 genomic construct without the upstream 1.4 kb promoter fragment and which thus ended 2.8 kb upstream of the PR-1 ATG (pAtPR1-R). This plasmid was used as a template in PCR with a left-to-right "topstrand" primer extending from positions -237 to -214 (DC39) upstream of the PR-1 ATG (oligo A) and a right-to-left "bottomstrand" primer comprising 15 bp of LUC coding sequence extending up to the LUC ATG and a further 19 bp of PR-1 sequence extending from the ATG into the PR-1 untranslated leader (oligo B: sequence: TTT GGC GTC TTC CAT TTT TCT AAG TTG ATA ATG G; SEQ ID No. 18). This PCR reaction was undertaken for five cycles at 94°C (30 s), 40°C (60 s), and 72°C (30 s) followed by 25 cycles at 94°C (30 s), 55°C (60 s) and 72°C (30 s) and this generated a product of 245 bp through annealing of the homologous PR-1 sequences; the fragment included a BglII site at its left end from the PR-1 promoter. A second PCR reaction was done using plasmid pDO432 as a template and using a left-to-right "topstrand" oligonucleotide which comprised 15 bp of PR-1 untranslated leader up to the PR-1 ATG and a further 12 bp of LUC sequence from the ATG into the LUC coding sequence (oligo C: sequence: TAT CAA CTT AGA AAA ATG GAA GAC GCC AAA; SEQ ID No 19) and a right-to-left "bottom strand" oligonucleotide extending from positions 332 to 312 (DC53) into the LUC coding sequence (oligo D). This PCR reaction was done under the same conditions as the one described above and generated a fragment of approximately 300 bp through annealing of the homologous LUC sequences; this fragment included a PstI site at its right end, derived from the LUC sequence amplified.

The two PCR fragments generated above were gel purified using standard procedures to remove oligonucleotides and were then themselves mixed in a further PCR reaction ("inside-outside PCR") with oligonucleotides A and D as primers. Conditions for this reaction were the same as described above. The amplified fragment was a fusion of the PR-1 promoter fragment from the first PCR reaction described above and the LUC 5' coding sequence from the second PCR reaction described above and had a BglII site at its left end and a PstI site at its right end. The fragment was gel purified and cleaved with BglII and PstI

to yield a product of 545 bp in size which was cloned into pAtPR1-R which had previously been cleaved with the same enzymes. Cleavage of the resultant plasmid (pAtPR1-S) with XbaI released a PR-1 promoter fragment extending from -2.8 kb to the XbaI site downstream of the LUC ATG, the fusion point between the PR-1 promoter and the LUC coding sequence being at the ATG. This fragment was cloned into XbaI cleaved pAtPR1-Q regenerating the full-length PR-1 promoter (4.2 kb) in operational fusion to LUC (pAtPR1-R).

**Example 4C: Transfer of the *Arabidopsis* PR-1 Promoter - Firefly Luciferase Gene Fusion to pCIB200**

TJS75Kan is first created by digestion of pTJS75 (Schmidhauser and Helinski, *J. Bacteriol.* 164: 446-455 (1985)) with NarI to excise the tetracycline gene, followed by insertion of an AccI fragment from pUC4K (Messing, J. and Vierra, J., *Gene* 19: 259-268 (1982)) carrying a NptI gene pCIB 200 is then made by ligating XbaI linkers to the EcoRV fragment of pCIB7 (containing the left and right T-DNA borders, a plant selectable nos/nptII chimeric gene and the pUC polylinker, Rothstein, S.J. et al., *Gene* 53: 153-161 (1987)) and cloning XbaI digested fragment into SalI digested TJS75Kan.

pAtPR1-R was cleaved with XbaI and SacI and transferred to SalI/SacI cleaved pCIB200 to create a binary vector construction (pAtPR1-S) suitable for *Arabidopsis* transformation pAtPR1-S was then transferred to *Agrobacterium tumefaciens* strain A136/pCIB542 for transfer to *Arabidopsis* Dijon-0 using the method described by Wen-jun and Forde, *Nucl. Acids Res.* 17: 8385-8386 (1989) T2 or T3 lines carrying the PR1-LUC transgene in the homozygous state were generated for chemical induction analysis.

**Example 4D: Chemical Induction of the *Arabidopsis* PR-1 Promoter**

Leaves of transgenic *Arabidopsis* lines carrying the PR-1 promoter-LUC gene fusion were treated by spraying with 0.15 µg/ml INA (isonicotinic acid) for 48 h. Five days later, sprayed leaves were analyzed for luciferase activity using the Promega Luciferase Assay System (Cat. # E 1500). Transgenic lines showed 144-fold induction over controls. In other assays, luciferase activity was measured by imaging lines for bioluminescence using intensified cameras (VIM) and photon-counting image processors (ARGUS-50 or ARGUS-100) from Hamamatsu Photonic Systems (Bridgewater, NJ). Leaves were sprayed with INA and subsequently with a 5 mM solution of D-luciferin (Analytical Bioluminescence

Laboratories, San Diego, CA) 24 h before and then immediately before imaging. Transgenic plants carrying the PR1-LUC fusion had strongly induced bioluminescence when compared to water-treated controls.

**Example 5: Isolation of the Maize PR-1 and Thaumatin-like cDNAs**

The maize PR-1-like cDNA PR-1mz was isolated by screening a BTH-induced cDNA library of maize with a probe matching to the PR-1 barley clone HVPR1BR (EMBL data base Accession # X74940). The DNA sequence of the gene and the deduced amino acid sequence are shown in SEQ ID NOs. 29 and 30, respectively.

The maize thaumatin-like cDNA PR-5mz was isolated by screening a BTH-induced cDNA library of maize with a probe matching to a known rice thaumatin clone (Reimann and Dudler, Plant Physiol. 101:1113-1114). The DNA sequence of the gene and the deduced amino acid sequence are shown in SEQ ID NOs. 31 and 32, respectively.

**Example 6A: Transformation of Disease Resistance Genes into Maize**

Ciba maize inbred line CGA00526 was transformed (Koziel, et al., Biotechnology 11:194-200, 1993) with tobacco *PR-1a*, tobacco *SAR 8.2*, and wheat *Thionin* disease resistance genes. Constructs used for transformation include the chimeric plasmid pUBA (Toki et al, Plant Physiol. 100: 1503-1507, 1992) containing the promoter, first exon, and first intron of the maize ubiquitin gene (Ubi-1), and the coding sequence of the *bar* gene conferring resistance to phosphinotrichin (Thompson et al., EMBO J. 6: 2519-2523, 1987). For cointegration purposes, the coding sequence of either the tobacco *PR1a* gene (Alexander et al, Proc. Natl. Acad. Sci. USA 90: 7327-7331, 1993), the tobacco *SAR8.2* gene (Alexander et al., 1992, Mol. Plant-Microbe Inter. 5:513-515), or the thionin gene from wheat (Rodriguez-Palenzuela et al., 1988, Gene 70:271-281), each under the control of said ubiquitin promoter were constructed into this plasmid resulting in three separate constructs for transformation. PCR techniques were used to obtain suitable restriction sites around the coding sequences of *PR1a*, *SAR8.2*, and wheat thionin, respectively. Maize tissue was bombarded with each plasmid separately or cotransformed with the *PR-1a* and *SAR8.2* constructs. Expression of the tobacco disease resistance genes *PR-1a* and *SAR 8.2*, a combination of *PR-1a* and *SAR 8.2*, and the wheat *Thionin* gene (Rodriguez-Palenzuela et al., Gene. 70:271-281, 1988) was confirmed by Northern blots.

**Example 6B: Transformation of Disease Resistance Genes into Wheat**

Immature embryos of spring wheat were isolated aseptically approximately 2 weeks after pollination, and plated scutellum side uppermost on MS based medium (Murashige and Skoog, 1962 *Physiol. Plant* 15:473-439) supplemented with 2,4-D, glutamine and asparagine. Four to six hours prior to transformation embryos were transferred to medium also containing 0.25M mannitol. Constructs used for transformation included the chimeric plasmid pUBA (Toki et al, *Plant Physiol.* 100:1503-1506, 1992) containing the promoter, first exon and first intron of the maize ubiquitin gene (Ubi-1), and the coding sequence for the bar gene conferring resistance to phosphinothricin (Thomson et al, *EMBO J.* 6:2519-2523, 1987). This construct was co-transformed with a single plasmid construct containing the coding sequences for the two disease resistance genes PR1a (Alexander et al, *Proc. Natl. Acad. Sci. USA* 90:7327-7331, 1993) and the SAR 8.2 (Alexander et al, *Mol. Plant-Microbe Inter.* 6:513-515, 1992) from tobacco, expression of both sequences being driven by the double 35S promoter. For cloning purposes PCR techniques were used to obtain suitable restriction sites around the coding sequences of PR1a and SAR 8.2, respectively. Plasmid DNA was precipitated onto gold particles as described in the Dupont Biostatic manual. Embryos were bombarded using the PDS 1000 biostatic device (Weeks et al, *Plant Physiol.* 102:1977-1084, 1993) and transferred the following day to the original medium without added mannitol. Calli which developed from the bombarded embryos were selected for a morphogenic response and regenerated on medium containing Basta. Presence of transgene fragments were confirmed by PCR. Total RNA was extracted (Lagrimini et al, 1987 *Plant Physiol.* 84:438-442) and gene expression confirmed by Northern analysis. Transformed plants contained and expressed the bar gene and the PR1a gene. Some events also expressed the SAR 8.2 gene. Plants were selfed and Basta resistant progeny obtained.

**Example 6C: Alternative Methods for Introducing and Expressing More Than One Anti-Pathogenic Sequence In Plant Tissue**

In addition to the possibility of expressing more than one transgene in transgenic lines by the sexual crossing of lines which are transgenic for one gene, the skilled artisan recognizes that an equivalent way of generating lines transgenic for more than one gene is by the use of transformation plasmids which carry more than one gene. For example, the expression of two cDNAs in a transgenic line can be achieved by the transformation of the host plant with a vector carrying each cDNA under the independent regulation of two promoters *i.e.* the vector carries two expression cassettes in addition to sequences needed for antibiotic selection *in vitro*. Each expression cassette can also carry any signal

sequence, vacuolar targeting sequence and transcriptional terminator so desired. Vectors carrying multiple expression cassettes can be constructed for use with *Agrobacterium* transformation or direct gene transfer transformation systems.

A further method for the expression of more than one transgene in a transgenic plant line is to firstly transform with a single gene (with appropriate regulatory signals) carried on a transformation vector and subsequently transform a line selected from this transformation with a further gene (with appropriate regulatory signals) carried on a different plasmid which utilizes a different antibiotic selection system. This method is obvious to those of skill in the art.

***Example 6D: Synergistic Effect of Combined Anti-Pathogenic Sequences***

Overexpression of two or more PR proteins in a transgenic plant gives rise to a synergistic anti-pathogen effect. The table below shows data from an experiment in which control tobacco lines and tobacco lines rendered transgenic for PR-1a, SAR8.2, or PR-1a and SAR 8.2 were inoculated with the pathogen *Peronospora tabacina* 5 and 8 days after inoculation, percentage leaf area infected was assessed. The constructs used to obtain the transgenic tobacco lines are similar to the constructs described in examples 6A and 6B. Tobacco lines transgenic for both PR-1a and SAR8.2 were obtained according to example 6C. The left side of the table shows raw data from numerous lines of each phenotype, whereas the right side shows mean data for each phenotype. In addition to the mean values for percentage leaf area infected, the relative area compared as a percentage of the control (=100%) is presented. This enables a calculation of the expected value in the line expressing both PR-1a and SAR8.2 assuming that the individual components are additive in action. The observed values of 40.8 and 30.4 % are well below the expected values of 63.4 % ( $0.756 \times 0.838 \times 100$ ) and 62.4 % ( $0.694 \times 0.899 \times 100$ ) for 5 and 8 days post-inoculation. These results demonstrate that the disease resistance effects of PR-1a and SAR8.2 are synergistic in transgenic plants.

Table A

	MEAN DATA					
	DAYS AFTER		DAYS AFTER			
	INOCULATION		INOCULATION			
	5	8			5	8
Control (L6)	36.1	42.2	CONTROL		40.2	49.4
Control (L18)	42.8	51.1		(100)	(100)	
Control (L22)	41.7	55.0				
PR-1a (L2)	35.0	42.8	PR-1a		30.4	34.3
PR-1a (L7)	25.0	25.0		(75.6)	(69.4)	
PR-1a (L13)	31.1	35.0				
SAR8.2 (L1)	36.7	49.4	SAR8.2		33.7	44.7
SAR8.2 (L8)	36.1	45.0		(83.8)	(89.9)	
SAR8.2 (L11)	26.7	45.0				
SAR8.2 (L21)	35.6	46.1				
SAR8.2 (L25)	33.3	37.8				
PR-1a/SAR8.2 (L17)	16.1	15.0	PR-1a/SAR8.2	16.1	15.0	
				(40.8)	(30.4)	
<u>Expected Additive Values:</u>				<u>63.4</u>	<u>62.4</u>	

***Example 7: Expression of SAR/CHX-Independent genes in transgenic plants***

The cDNAs described in examples 1A and 1B can be expressed in transgenic plants using techniques well known in the art. As components of the signal transduction pathway involved in SAR, the CHX-independent genes are useful for the manipulation of the SAR response. For example, the constitutive expression of key components in the SAR transduction pathway in transgenic plants will likely lead to the generation of plants with

enhanced disease resistance characteristics and this will likely be achieved by the activation of components in the pathway downstream to the component being expressed transgenically and hence to the activation of anti-pathogenic end products. By way of illustration this may be achieved from the expression of the appropriate genes behind the constitutive 35S promoter cDNAs may be transferred to the vector pCGN1761 or pCGN1761/ENX which carry the double 35S CaMV promoter and the *tm*/transcriptional terminator on a pUC-derived plasmid. Colonies carrying the cDNA in sense are recovered and the cDNA carrying expression cassette is subsequently excised and cloned into pCIB200 for use in plant transformations using *Agrobacterium*. For direct gene transfer, the cDNA-carrying expression cassette is transferred to the vector pCIB3064. Transformation to transgenic plants is undertaken using techniques well known in the art. For transformation of dicotyledonous species using binary *Agrobacterium* vectors such as pCIB200 see Alexander *et al.*, *Proc. Natl. Acad. Sci.* 90: 7327-7331 (1993), and for transformation of monocotyledonous species using direct gene transfer vectors such as pCIB3064 see Koziel *et al.*, *Biotechnology* 11: 194-200 (1993). Transgenic plants are screened for high-level expression of the appropriate cDNA by Northern or Western analysis. Plants which express high levels of the gene product are found to have enhanced resistance to plant pathogens.

Other promoters are suitable for the expression of these cDNAs in transgenic plants. These include (but are not restricted to) constitutive promoters (such as those from the ubiquitin and actin genes) and cell and tissue-specific promoters.

For genes involved in the signal transduction of SAR which may cause negative regulation of the SAR pathway, increased disease resistance can be achieved from the constitutive expression of cDNA in antisense to the gene coding sequence. The cloning and transfer of antisense sequences is undertaken in the same way as described above, except that the orientation of the cDNA is inverted to effect expression of antisense transcripts.

\*\*\*\*\*

While the present invention has been described with reference to specific embodiments thereof, it will be appreciated that numerous variations, modifications, and embodiments are possible, and accordingly, all such variations, modifications and embodiments are to be regarded as being within the spirit and scope of the present invention.

LIST OF REFERENCES

Alexander *et al*, *Proc. Natl. Acad. Sci.* 90: 7327-7331 (1993)

Alexander *et al*., Molecular Plant-Microbe Interactions 5:513-515, (1992)

Allen, G., *Plant Sci. Lett.* 26: 173 (1982)

Altschul *et al*, *J. Mol. Biol.* 215: 403-410 (1990)

Ausubel *et al*, in "Current Protocols in Molecular Biology, pub. by J. Wiley & Sons, New York (1987)

Berk, A.J. *et al*, *Cell* 12, 721 (1977)

Bevan, M.W. and Chilton, M-D., *Int. Rev. Genet.* 16: 357 (1982)

Boulter *et al*, *Crop Protection* 9: 351-354 (1990)

Brugidou *et al*, *Mol. Gen. Genet.* 238: 285-293 (1993)

Carozzi *et al*, *Plant Mol. Biol.* 20: 539-548 (1992)

Chester, K.S., *Quart. Rev. Biol.* 8: 275-324 (1933)

Church and Gilbert, *Proc. Natl. Acad. Sci.* 81: 1991-1995 (1984)

Claes *et al*, *Plant Cell* 2: 19-27 (1990)

Davis, *et al*, *Proc. Natl. Acad. Sci, USA* 81: 2194-2198 (1984)

De Pater and Schilperoort, *Plant Mol. Biol.* 18: 161-164 (1992)

DeBlock, M. *et al*, *EMBO J.* 6: 2513 (1987)

Deveraux *et al*, *Nucl. Acids Res.* 12: 387-395 (1984)

Evans, D.A. *et al*, in: "Handbook of Plant Cell Culture", Vol. 1, MacMillan Publ. Co., 1983, p. 124

Fischhoff *et al*, *Biotechnology* 5: 807-813 (1987)

Greenberg *et al*, *Mol. Cell Biol.* 6: 1050-1057 (1986)

Gubler, U. and Hoffman, B. J., *Gene* 25: 263 (1983)

Hanahan, D., *J. Mol. Biol.* 166: 557-580 (1983)

Hilder *et al*, *Nature* 330: 160-163 (1987)

Klein, T.M. *et al*, *Nature* 327: 70 (1987)

Koziel *et al*, *Biotechnology* 11: 194-200 (1993)

Kuc, J., *BioScience* 32: 854-860 (1982)

Lagrimini *et al*, *Proc. Natl. Acad. Sci.* 84: 7542-7546 (1987)

Lau and Nathans, *Proc. Natl. Acad. Sci.* 84: 1182-1186 (1987)

Liang and Pardee, *Science* 257: 967-971 (1992)

Margis-Pinheiro *et al*, *Plant Mol. Biol.* 17: 243-253 (1991)

Mauch *et al*, *Plant Physiol.* 76: 607-611 (1984)

Meins *et al*, *Mol. Gen. Genet.* 232: 460-469 (1992)

- 50 -

Messing, J. and Vierra, J., *Gene* 19: 259-268 (1982)  
Mikkelsen *et al*, in "Advances in Chitin and Chitosan, ed. by Brine *et al*, pub. by Elsevier, Amsterdam (1992)  
Murashige and Skoog, 1962 *Physiol. Plant* 15:473-439  
Ow *et al*, 1986; *Science* 234: 856  
Paszkowski, J. *et al*, *EMBO J.* 3: 2717 (1984)  
Potrykus, I. *et al*, *Mol. Gen. Genet.* 199: 169 (1985)  
Reich, T.J. *et al*, *Bio/Technology* 4: 1001 (1986)  
Reimann and Dudler, *Plant Physiol.* 101:1113-1114  
Rodriguez-Palenzuela *et al.*, *Gene*. 70:271-281, (1988)  
Ross, A.F., *Virology* 14: 329-339 (1961)  
Ross, A.F., *Virology* 14: 340-358 (1961)  
Rothstein, S.J. *et al*, *Gene* 53: 153-161 (1987)  
Saiki *et al*, *Science* 239:487-491 (1988)  
Schmidhauser and Helinski, *J. Bacteriol.* 164: 446-455 (1985)  
Simpson, R.J. *et al*, *Biochem. Intl.* 8: 787 (1984)  
St.John and Davis, *Cell* 16:443-452 (1979)  
Stahl, *et al*, *Nuc. Acids. Res.* 16: 3026-3038 (1988)  
Thompson *et al.*, *EMBO J.* 6: 2519-2523, (1987)  
Toki *et al*, *Plant Physiol.* 100: 1503-1507 (1992)  
Uknes *et al*, *Plant Cell* 4: 645-656 (1992)  
Uknes *et al*, *Plant Cell* 5: 159-169 (1993)  
Vaeck *et al*, *Nature* 328: 33-37 (1987)  
Van Loon, L.C. *et al*, *Neth. J. Plant. Path.* 88: 237-256 (1982)  
Ward, E. *et al*, *Plant Cell* 3: 1085-1094 (1991)  
Weeks *et al*, *Plant Physiol* 102:1977-1084, 1993  
Wen-jun and Forde, *Nucl. Acids Res.* 17: 8385-8386 (1989)  
White, R.F.; *Virology* 99: 410-412 (1979)  
Williams *et al*, *Bio/Technology* 10: 540-543 (1992)

EP-A 223 452

U.S. Patent No. 4,751,081

WO 87/07299

EP-A 240 332

U.S. Patent No. 4,795,855

EP-A 249432

U.S. Patent No. 4,940,840

EP-A 392 225

**DEPOSIT**

NRRL B-21169

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: Ciba-Geigy AG  
 (B) STREET: Klybeckstr. 141  
 (C) CITY: Basle  
 (E) COUNTRY: Switzerland  
 (F) POSTAL CODE (ZIP): 4002  
 (G) TELEPHONE: +41 61 69 11 11  
 (H) TELEFAX: +41 61 696 79 76  
 (I) TELEX: 962 991

(ii) TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC DNA SEQUENCES AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 29

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1079 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ACAAACCAAT AACACAAAAC ATGGCTTC CAATAATCTC CTTAGTCCTT CTTCTCTGCC	60
TCTTAGGTTT CTTTTCTGAA ACTGTCAAGT CTCAAAACTG CGGTTGGCTT CCAAAACCTCT	120
GTTGCAGTCA GTTCGGTTAC TGTGGTACCG ACGATGCATA CTGCGGGTGT GGATGCCGAT	180
CAGGTCTTG TAGAGGTAGT GGAACCCCGA CCGGAGGGTC GGTOGGTAGC ATTGTGACAC	240
AAGGTTCTT TAACAATATT ATCAACCAAG CTGGTAATGG TTGCGCGGGG AAAAGATCT	300
ACACCCGTGA CTCTTTCTT AACGCCGCTA ATACCTTCCC CAACTTTGCC AATTCTGTTA	360
CCAGACGTGA AATTGCTACC ATGTTTGCTC ATTTCACTCA CGAGACCGGA CATTCTGCT	420
ACATAGAAGA GATTAACCGA GCAACACGTA ACTACTGCCA GAGCAGCAAC ACACAATACC	480

CATGTGCACC GGGAAAAGGC TACTTCGGTC GTGGTCGAT CCAACTATCA TGGAAC TACA	540
ACTACGGAGC GTGTGGTCAA AGTCTCGGTC TTGACCTTCT ACGCCAGCCC GAACTTGTGG	600
GTAGCAACCC AACTGTAGCT TTCAGGTCGG GTTTGTGGTT TTGGATGAAT AGCGTAAGGC	660
CGGTTCTGAA CCAAGGGTTT GGAGCCACCA TTAGAGCTAT TAATGGAATG GAATGTAACG	720
GTGGTAATTG CGGTGCAGTC AACGCAAGGA TTGGATACTA TAGAGACTAT TGTGGACAGC	780
TTGGTGTGGA CCCTGGTCCT AACCTTAGTT GCTAAAAAAC CTTTGAACCC AAACACGGAC	840
ATATGTGACG TGGCATGTAA TAAGTGAGAT ATACTAAAAT TTACACACGTA TGTACTTTAT	900
GTCGGGTCTC GGTGTTCCCT GCGTCACAAG CAAAAAATTG TTGTAATAAA CTTGTGAAAG	960
TGATTTCTT TTCTTATGTG ACTTCCTTATG AAAGAGAATT TTAAGATTGA TAGATGTTTG	1020
AGTTTGACAC TTGCA GTATC TGATTTGGG ATACTCGCAT AAAAAAAA AAAAAAAA	1079

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 952 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

AAAACAACAT AACATGGCTT CCACCAAAAT TTCCCTTAGTC TTTTTCCCTT GCCTCGTAGG	60
TCCTTGTATA GGGGCTGGAA CCACAAAGCGC TGOOGAATCA GTCACTAGCA TTGTGACACA	120
AGGTTTCTTT AACAAACATTA TCAACCAAGC CGGTAAATGGC TGCACGAGAA AAAGATTCAA	180
CACCCCGCGAC TCCTTCATTG ACGCTGCTAA TAATTTCCA AACTTTGCAA GTTCCGTTAC	240
AAGACGTGAA ATTGCAAGCA TGTGGCTCA TGTCACTCAC GAGACCGGAC ACTTCTGCTA	300
CATAGAAGAA ATAAAGAAAA GGTCACTGGG TAGGTGCGAC GAGAACGTAG AGCAGAGACC	360
ATGTCCCATCA CAGAGTAAAG GTCACTCCGG TCGTGGTCAT CAGAGTCCTG GTCTCGACCT	420
TCTACGCCAG CCTGAGCTTG TGGGTAGCAA CCCAACTGTA GCTTTCAAGGA AGGGTTTGTC	480
GTTTTGGATT AATAGCGTAA GGCGGGTGCT GAACCAAGGA TTTGGAGCCA CTATAAGAGC	540
CATCAATGGG ATGGAATGTA ACGGTGGGAA TTCAGGTGCG GTCAAGGCCA GGATTGGGTA	600
CTATAGAGAC TATGTAGAC AGCTTGGTGT GGATCCCTGGT CATAACCTTA GTTGTAAAAA	660

CGTTTTACAA ACGCATACAA CGTGACGTGT CACGTGATAA TGCGGAATAA GAAGTTACAT	720
TTCAAAACAGG GCAAGGATTG AGAATCGCAA GTACATGAAT TAACATTCTG TAACTGTTAA	780
TTGTATCAAG TTGTGTTAAT ATATAAGAGC TAGGAGAATT GGTCTTCTT ATGTCTAGCT	840
AGTAACCATC TATGTTGTA GCTAGGAGCT ACTTGTGTT TCTAGTGTGTT TGCTATCTA	900
AGCCAATCAT ACAAAATCTA TGCTTATTCA CAAAAAAA AAAAAAAA AA	952

## (2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1250 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTGGCA CGAGCACACA CAATCCGCC TATAGCCACC ATGTCTCTC AGCCTACCTT	60
AGCCACCACC ACCACGCGTG CTACTGGGCC TAAACCTTCC TATTACAAA CAGCTCTGT	120
TAGCCAAGAG AGTATCCAGC GAACGGAGGT TCTTCCTCAC CTGTACGCCT ACCAGCACAC	180
CCAAGGAAAA CCAAATGCTA ACCAGACTGT CATACTGGAT CCAAAGCTCC CGCGTGTGTT	240
CGGTGGCCCTT GCTGCTAATG ACTGGACCAT ATATGATGGC ACTTGCACCA ATGCAAACCT	300
TGTTGCGCAC GCTCAAGGTT TGCACATTGG TGCTGGTATG ACCAAAGAAA ATTGGTTCAT	360
TTGTTTCAAC ATGGTGTGTTG TGGACCGAAG GTTATGGGT TCCAGCTTCA AGGTGATGGG	420
AGATTTTCGA GGAAACGAAG GTGAATGGGC AATTGTTGGT GGGACGGGAG AGTTTGCATA	480
TGCACAAGGT GTCATAACCT TTAACAAGAC CTGGTOGCC CAGGCAAATG TCAGGGAGCT	540
TCATGTTCGT GCTTTGTGTC TGTCCCTCTC AAAAGCACCG GAAACACCGT GCTCAAGGAC	600
ACCAACGGGAG AGCTCTGTCA CCAAGATCGG CCCATGGGT AAAATAAGTG GAGAATTCT	660
TGACGTTCCC ACGACACCCAC AACGTCTAGA GTGTGTGACC ATTGCCATG GAGTTGTCAT	720
TGATTCACTT GCATTTCTC TCGTCGACCA AGCTGGTGG CAACATAACG TTGGCCATG	780
GGGTGGGCCA TCGGGGGACA ACAAGGACAC GATTAACCTT GGTCCATCGG AGATTGTGAC	840
AGAAGTCTCT GGAACGATTG GTGTATTGG AGCAGGCAAT GTCGAGTACA ATGCCATAAC	900
ATCACTAACCA ATTACCAACAA ATGTCGCGCAC GTACGGGCGC TTGGAGAAC CGCAGTGTAC	960

- 54 -

TCGTTTCAGT GTTCCCGTGC AGGACAAAAG CAGCATCGTG GGTTTCTTCG TGTGCGCTAG	1020
GAAATACGTG GAGGCGCTCG GGGTTTACGT GTGTCACCT ATTCAAACCT AGTCCCGAAG	1080
TGCTTCACAC TCGGTGTGCC TTCCACCTTT CTATGTTGTG CCAATAAAAGT AGGTTATATA	1140
CTTTGGTGT AATGCTATGC CCCCCCCCCAA GTGTGATGGT TTTCATATCA GCAAGAAGTT	1200
TCTCTTGCTA CTATATGCAA AATAAAAAAA AAAAAAAA AAAACTCGAG	1250

## (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 340 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGCTCGNGTG GCGAACGGAC GAGGAGTTCG CCCGGAGAT GCTCGCCGGN CTCAAACCCAC	60
ACATCATCAC CGCCTGAAAC GTGTTCCCTC CGAGGAGCAC GNTTGAAGGG TACGGTGACC	120
AGACGAGCAA GATCACGGTG GAGCACATCC AACACAAACCT CGGNAAGCTC AACGTCGACA	180
AGGCAATOGA CGCCAAGAGG CTCTTCATCC TAGACCACCA CGACAACCTC ATGCCTCACC	240
TGCTAAAGAT CAACAGTCTC CCAAACACAT TCCGTCTTAC GNCACCAGGA CGCTGCTCTT	300
CCTGCAAGAT GACGGGACTC TCAAGNCAAT CGNCATCGAG	340

## (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 436 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ACATCGAGAA GGGATCATGA ATTGACCCGA TCCCCAACTC AAGAACCGGA ATGGGCCACC	60
AAAAATCCCCCT AATGCTACTC TACCCCAACA CATCTGACAT CGATGGTGAG AGCGCCACCG	120

- 55 -

GGATCACAGC CAAGGGCATC CCCAACAGCA TCTCCATCTG ATCGATGTTA TGCCGTTTA	180
TGTTATGTT GTATGTTCC CATTGCAAAT AAGGAACGCT GCATGTGCAC GTTTCATGAG	240
TGGCCAGAAG CGTGCTCGTT CACGTTGAGG CGAGTTGTGT TTATTTCTGT GTTGATGAAG	300
TGTTTTGGCC GCAATGGAAA GCCC GTGGGA TGGCACGGAT GAAGGGTTGT GGTCGGCCAG	360
TGTGTTGAGC TGGACATATA TACGGAAAAA AATTGTATGA GACCGGGTCT CATAGACCCA	420
AAAAAAAAAA AAAAAA	436

## (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1162 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AAGACACAAC CTTATTTGC ATATAGTAGC AAGAGAAACT TCTTGCTGAT ATGAAAACCA	60
TCACACTTGG GGGGGGGCA TAGCATTACA CCAAAAGTAT ATAACCTACT TTATGGCAC	120
AACATAGAAA GGTGGAAGGC ACACCGAGTG TGAAGCAGTT CGGGACTAGT TTGAAATAGG	180
TGGACACACG TAAACCCCGA GGGCCTCCAC GTATTCCTAG CGCACACGAA GAAACCCACG	240
ATGCTGCTTT TGTCTGCAC GGGAACACTG AAACGAGTAC ACTGCGGTTC TCCAAAGGGC	300
CCGTACGTGC GGACATTGTG GGTAAATGGTT AGTGTATGTTA TGGCATTGTA CTOGACATTG	360
GCTGCTCAA ATACACCAAAT CGTTCAGAG ACTTCTGTCA CAATCTCCGA TGGACCAAGT	420
TTAATCGTGT CCTTGTGTC CCGGCATGGC CCACCCCATG GGTCAACAGC CCACAGCTTC	480
ACAAGCCAAT ACACACACCA ATATAGTCTC CTTTCCAGAG CTAGCTAGCA TCAATGGCTT	540
CTGGTGCTCA GATGAAGATG GTTGCCTGCG GCATGATGCT GGCCATCATG TTCATCGCTG	600
CAGCTCATGC AGAACCCAGGG CCTGCAGAAA CTTGCATCGA CAAGACCGAA AAAGTTGGTC	660
TTGCCACTGA CTGCATCTGC TCCAAGAACT GTGCTTGTGC AGGAAAGTGC ATCTTAAAG	720
GGGGCGAGGG TGGCGAAATC CAGAAGTGCT TTGTCGAATG CGTGCTGAAA AATGACTGCA	780
ACTGCAATGC TAAGCACCAAC AGTGCCTGCAG CCCCTCAGTA AGGAGACTCC TGATTGGAGC	840
TTCTTCAGCC ATGCTATGAG GTCTGCGTCC ACAAGTCCAC AGTCAACATT TGAGTAAATT	900

AATAAACCTC TTATACCTGC GTTCTTGGCA TGGCCTTATT GCTTGCAGCG AATGTATGTT	960
GAGAGTTGGT TTCCCGGTCA CAGCTCATAT ATATGCTGCC AGTCACGTGT ACGACGACTG	1020
CCACATGTAT TGCTGGAAAT GAATATGTAT GTGTATTGT CCTTTTCGTC ATTTGGGTCA	1080
TTTTAAAAAT AAGTTGTATG CTGGAGCATG ACAACAAGTG TATTGGCTA CTATAATATT	1140
ATTAATGATA TGATGTCCTG TT	1162

## (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1371 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATCCACCTCG ATCCAGCCAT GCTGTCCTCC AAGTCCTCCT TAGCTGTTGC TCTAATCCCT	60
GTGGTCACCC TGACAGATGT GCTACCGCTG GTATCTTCCT CACGTGGCCT CGTTGGGTAA	120
CCCACTGATG GGGCACTGGA GGACTCTCTG TTGATGATGG AGAGATTCCA TGGCTGGATG	180
GCAAAGCATG GCAAGTCGTA TGCGGGAGTC GAGGAGAAAC TGCGGCGGTT TGACATATTC	240
CGCAGGAACG TAGAGTTCAT CGAGGCGGCG AACCGAGATG GCAGGCTCTC GTACACCCCTC	300
GGGGTGAACC AGTTCGCCGA CCTCAOCCAC GAGGAGTTCC TTGCCACGCA CACCAGCCGC	360
CGTGTGGTGC CGTCAGAGGA GATGGTGATT ACAACTCGCG CTGGCGTTGT TGTCGAGGGT	420
GCCAAATTGTC AGCGGGCGCC AAATGCTGTG CCTCGTAGCA TCAATTGGGT GAATCAAAGC	480
AAAGTCACCC CAGTCAAAAA TCAAGGAAAA GTATGCGGGG CTTGCTGGGC TTTTTCTGCC	540
GTGGCCACGA TCGAAAGCGC CTACGCGGTC GCGAAGCGAG CGAGGCCGCC GGTTCTGTC	600
GAGCAGGAGC TCATCGACTG TGACACAATC GACAGAGGCT GCACGAGCGG CGAGATGTAC	660
AATGCCTACT TCTGGGTCTT GAGGAACGGC GGCATCGCCA ACAGCTCAAC GTACCCCTAC	720
AAAGAGACTG ACGGCAAGTG CGAGAGAGGG AAACGCGAGG AACACGCGGC CACGATCAGG	780
GACTACAAAT TCGTCAAACC CAACTGCGAG GAGAAGCTCA TGGCAGCCGT GGCGGTGOGA	840
CCCGTOGCGCG TCGGGTTCGA CTCCAACGAC GAATGCTTCA AGTTCTACCA AGCTGGTTTG	900

- 57 -

TACGACGGCA TGTGCATCAT GCACGGGAA TACTTTGCC CGTGCTCGTC CAACGACCGC	960
ATCCACTCCT TGGCCATTGT CGGGTACGCC GGCAAGGGGG GCGACAGGGT CAAGTACTGG	1020
ATCGCCAAGA ACTCGTGGGG CGAGAAGTGG GGAAAGAAGG GCTACGTCTG GCTGAAGAAG	1080
GATGTTGATG AGCCGGAAGG CCTCTGCCGC CTTGCAATTG AGCGGGTATA TCCTATAGTC	1140
TGATCTGATC TGACGAGATC GACTGCAGTG GGCGTGCATG AAACCTACGG AAATGGCATT	1200
CACCTATATT TTGGGTTGCT CTGTATGCAT GGATGCGCT ACTATATTTT ACTACATATA	1260
TATTCATCTC CCGCTAATAA AACTACATGT CCTTGTATCC ATTTATGCAC GTTTATCCAT	1320
ATCTTGAATA AATTGGATGG ATTGGTTATC CAAAAAAA AAAAAAAA A	1371

## (2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 723 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CAACCGGGAA CCGGAGGCCT CGCACCCCTC ACCTCGCTGT GGTTGACAAC ATCGACCTGT	60
GCCCAAGACC TCGGGGAACG CTACCAACTC CGAACCCAAG TATGTOCCCCG CCTTGCACGA	120
GGCGGACGTC AAAGAAGGAG GTGAAGGGCG TTCTGTACCC ACGAGGGGAC GCACGTGTGT	180
CAGTGGAACG GGCAGGOCAG GTCAAACGGC GGGCTCATCG AGGGGATCGC CGACTACGTG	240
CGGCTCAAGG CCGACCTCGC GCGACGCAC TGGGOGCCGC AGGGGAGOGG CGACCGTTGG	300
GACGAGGGGT ACGACGTGAC GGCCAAGTTC CTGGACTACT GCGACTCCCT CAAGGCCGGG	360
TTCTGTGCGG AGATGAACAG CAAGCTCAAG GACGGATACA GCGAOGACTA CTTOGTGCAG	420
ATCCCTGGGA AGAGCGTGG CCAGCTGTGG AACGACTACA AGGCCAAGTA CCCCCCAGCC	480
CCAGGGCTGA TCGACGATGC ATGCAGTTTG TTGTGTATG TGTACCGGTC TTCTGTACACA	540
TACAGATACA TTATAGTACT TGTATTACTG TACAATTTAT GTACTGCCTG GAATGGAATA	600
AATCAGCGTT GGCACGGTGT GTGTTAACGA ATTGACGAGA CAAAGGGACC GTCTATAGGT	660
CATGTCTATCG GTTGCCTGAA ATACATTGAA CATCATCACT TTCTTTACAG CAAAAAAA	720
AAA	723

- 58 -

## (2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 765 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AAATGACAAT	TGAAATTAT	TAAAGTAATA	TGATCTAATA	TGTTCAAACA	AGCTCAATCA	60
CAATTAAATT	TAATTAATAT	TGACCGCGCA	TOGCCGAGT	ACGACTACTA	GTCAATTAA	120
GAAAGCTGGA	GCCATTGCG	GAGTAATCAT	GAAATTGGCG	GAGTTGACTT	COGATTCCGG	180
TGATGGTGCG	AGTATTCCCG	AGCTTTGAT	CTCCGCCATT	CCACGGGOTG	CATTTOCACC	240
TTGTTCTCA	CACAATCTGG	GAAGGTATAC	ACCTTCTCCA	GCAGCAAGGT	TGAAGTAGAG	300
ATCAATAATA	TTAGGGCAAT	TTTGTAGCA	CTGAGGGGAG	CACAGCTTCT	GGGTAAAGTG	360
GCACTCGAGG	AGAGAATCAG	AAGAATTCCG	AAAGTATTCC	TATCAACACC	GCAAGCTGAA	420
ATACACTGGT	CGGTTCAAT	CCAGTCCTTG	AGCTTATCAG	CCTCAATTTC	CGACGTCTTA	480
CATGTATATA	CTTCITCACC	ACTCCTTGG	AGACGTTCT	CTAACACGCA	ACGCTTCCCA	540
GTACTTGAAA	TCGAAAAGC	GCATGAGTCC	TTGTTTAGAT	TCTCGCATGT	TATGCTCCCT	600
AGAGTGACTT	GAACACAGAA	GACAAGTGCA	CAAGCAACAA	TAGCCAAAGT	AAAGTTGTGG	660
AATGAAGCAA	CCATGGTGAA	AAATCTAGCA	ACAATTGATC	AGACTATAAA	TTTTTCTGA	720
GTATATATAG	CTCCCTTGG	CTTCGACTC	TCCTTTTTG	TGGTT		765

## (2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1054 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

GAATTCTGGCA CGAGTTTGCG TATACTCAA TACTCCGGCT AAATCGGCAA ATATATCCAA	60
CAAATTAAGC CTTTTAAACC AATCAGTACG TTATCAATAT CATCAATGTC AGATCGTATT	120
GATGTCCTGG CTATGATGAT ATCAGATGAG GAATTGCGAG AGAAACTGCA GATGCAACAT	180
GTTCTCGGAG CCTCTTATGA AAGTCCCAA GGATTCTGTG GAATTGCAA TAAGAAAAAG	240
GGAGCAGATC GAATGTTCAA ACTCCAAATC TGCTCTGTC ATTCTTCTG TTCTCACTGC	300
ATAGGTCTTT ATGCCAAATC CAAGAATCGC GACAACGTTT TCCCTATAAC TTGTCCAGGT	360
TTGAGATGTC GTGTCGTAAT TCAACCGAA TCTTGTAGGT CTATCATTCC AGTAAATACA	420
TTTGCAGGT GGGAGAGAGGG CTCGGCCGAG TCAGCTATTTC CTAATGGTGA AAAATTCTAT	480
TGTCCTTATA AAAGATGTTTC AGGGCTGTTTC ATTCAATGATC GACATGAAGA GATCGTTCA	540
TGCATTTGTC CTCTGTCGAA AAAACTGTTTC TGTCGGAAGT GCGAGTCCTC TTGGCACACT	600
GGCGTGTGATT GTGACAAGTT TCAAAAGGTA GAAAAGACA GAGAAGATGA GCTAAAACCTT	660
AAGCTGCTTG CTGAGATCAA AAAATGGAAG GAATGCCCTA AGTGCCTAAATC CATTGTTGGAG	720
AAGGTTGATG GCTGTATACA CATGACTTGC AGGTGTAAAA TGGAATTTTG CTACGTATGT	780
GGAGGAACGT GGAGCGAAAG GCAATTGGAGT TGCCAAAGACT AGTTATGATA TACTATCGGA	840
TTGCCTCTTC ATACCAATCG GATGCTGAAT ATTTTGTGT CAGAATAAAA ATTATGTTGT	900
TTAGTTGGTT GTATTATGTC TTTAATTGG TGTCATGGT TTGTATTAGG ACCATTATAT	960
ATAGTACGAG TTGGTCCCTTG CATAACATTT ACTTGATGTT GATTATAATA TGCTCGAGTT	1020
GGAGACCCCA AAAAAAAA AAAAAAAACG CGAG	1054

## (2) INFORMATION FOR SEQ ID NO: 11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 760 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTCGTGCCTGCTTAAAGTGTG CTCGAACCTTC CTACTGGTAA TGAGGATGAT GATGACCTGG	60
AATTTGAAAAA TGTCCACTGG AATGGTTCAG ATATGGCCTTC CGATGATACT CAAAAATCTC	120

- 60 -

ATAGACCAAG GCACCGCGTA CATAAATCGT CTGGTTCCCA CAAGGCCATG AGCCGCTCCC	180
TTTCATGTGA CTCGCAATCA AAAGGATCTA TTTCTACACC TCGTGGGTCC ATGGTTGACC	240
TAAGCCTAACT CGAGATGGCT GCACGTGCGA GATATTGGCG ACACCTTAAC CTTAGGGAAG	300
CTATACCTAA CCCCTCGAAA GAGCAACTAA TTGATGTTGT CCAGAAGCAT TTACATCTC	360
AGCAATTGGA CGAGTTGCAG GTAATTGTGG GATTGTTCA AGCTGCCAAG AGACTCAAGA	420
CAGTCTGCAA ATGACTCAAG GGAACGAACC CCTTTGGGA ACGCCCTAA CACTAACAAA	480
TAGTGGTACC TTTGTGTCCA CTTGCTATTG CGCAGTAATA TGTGTAGTAT AATATGTATA	540
CTCCTGTTA TCCCTTATGT CTTGCTTACA GTTGTAGTTT ATGTAGATAA ATGAACTGCT	600
GAACTAACTC TGGAGGTCTC CTTCTGGTTG AATGTACTTC CTGCGAGTG GGGGAAGTAC	660
TAGTACTTAA TCACATGTAT AGTCTAATGT CTTAAACTGT TAAATCATCA TTCAACCAAC	720
TGAATCTTGG TTGGTATTA GTGACCCCTCG TGCGGAATT	760

## (2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 396 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TGCCGCTTTA AGTGTGCTCG AACCTCCCTAC TGGTAATGAG GATGATGATG ACCTGGAAATT	60
TGAAAATGTC CACTGGAATG GTTCAGATAT GGCATCOGAT GATACTCAAA AATCTCATAG	120
ACCAAGGCAC CGCGTACATA AATCGTCTGG TTCCCAACAAG GNCATGAGCC GCTCCCTTTC	180
ATGTGACTCG CAATCAAAAG GATCTATTTTACACCTCGT GGGTCCATGG TTGACCTAAG	240
CAAACTCGAG ATGGCTGCAC TGTGGAGATA TTGGGAGACAC TTAAACCTTA GGGAGGTATT	300
CCTAACCCCTC GAAAGAGCAC TTATTGATGT GGTCAAGAGCA TTCAANATCTC AGAAATGGC	360
GAGTGGAGGT ATTGTGGATT GGTCAAGCTG CAAGAG	396

## (2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 653 base pairs
- (B) TYPE: nucleic acid

- 61 -

(C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CTTGTGTTGGT TGTGAGTT ATTGCTTC TAAGAACATT GTGAGAAATG GCTGCTAACG	60
ATGCTACTTC ATCCGAGGAG GGACAAGTGT TCGGCTGCCA CAAGGTTGAG GAATGGAACG	120
AGTACTTCAA GAAAGGCGTT GAGACTAAGA AACTGGTGGT GGTGATTTT ACTGCTTCAT	180
GGTGGGGSCC TTGCGGTTTT ATTGCCCAA TTCTTGCTGA CATTGCTAAG AAGATGCC	240
ATGTTATATT CCTCAAGGTT GATGTTGATG AACTGAAGAC TGTTTCAGCG GGAATGGAGT	300
GTGGAGGCAA TGCCAACATT TGTCTTCATT AAAGATGGAA AAGAAGTGGAA CAGAGTTGTT	360
GGTGCCAAGA AAGAGGAGTT GCAGCAGACC ATAGTGAAGC ATGCTGCTCC TGCTACTGTC	420
ACTGCTTGAA TCTCCTTAAT CAAGGGGATG ATATCCATA TTTAGTAGTA TTGCTTTTG	480
TAATAACCAA GTAACTTGTG CGAATTTCAC ACTATGGATC ACTGTATGGT TGTACTATCC	540
ACCATGTTTT TATTGCTTTT GTGAACCCTG TCTTGTGCT TGGAAATCTGA TTTGTGCATT	600
ACTGGTGTAA GGCTATATGC CCAATTCTYAC AAAAAGACTA CTTTAGATT TCT	653

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1697 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATTCAAAATT ATTGGATCCT AGTTTAAAGT GGAACCTCTA ATTAAATCAAT CATTTCATGA	60
TCAACCTTTA CTAACATCAC CATTTCCTT GACAACATTA TTTCCGACT TACCATTTTC	120
CCTAACACCA TTACCCCTCTG GATTTCCCTT GACAATATTA CTCTCCGACT TCTCATTTC	180
GTTCACTTTA TCAGTAATCT TGGCTTCCTC AGCAGCTGCT TCTTCTCTT TCTCATTTC	240
TTTCTCAGCC CTTTCTCTTT CTTCTCCTC AGCCCTCAAT TTTGGCTCCT CCTTTGCACT	300

- 62 -

TTCAAGAGCT	TTAACAGTC	TCTCCAAACA	AGTGTGCA	TTTCCTTAG	AAGACTTGGG	360
CATTAATTC	TCAGCAACAT	CAGCAGGAGT	AAATATTAGT	TCCCCCAATA	AAIGACCAAT	420
CTCAGGAAAA	TGATCATGAG	ACTCAATATC	TAGATAATT	TTGCAAGCA	CCCTGAAGGA	480
GTCAAAGCAA	CAGTATGATA	ACACAATGTG	TTTATCCATC	CTCCCCCTCC	GAATTAAAGC	540
AGGGTCAAGC	TTTTCCACAA	AGTTGGTAGT	GAAAACAATA	AGCCTTCAC	CACCAATAGC	600
TGACCATAAC	CCATCAATAA	AGTCAAAAG	CCCAGATAAA	GTCACCTCGC	TTTGCTTTT	660
CTCCTCTCGA	TTTTTCATCT	TCTCCTTGAC	GGCATCTTTC	TCGTCTTTA	CTTCCTCTTC	720
CTTGTGCTCT	TTCTCTTCT	CCCTTTGGCC	GGTAAGGTCA	AGCGAACAGT	CGATGTCCTC	780
AATCACAATG	ATAGACTTAC	TAGTAGTATC	TATTAATAAC	TTTCTTAGCT	CGGTGTTGTC	840
CTTAACCGCT	GTCAATTCAA	GATCATAGAC	ATCATATTGT	AAGAAGTTAG	CCATTGCAGC	900
AATCATGCTA	GACTTACCGG	TTCTGGAGG	ACCATATAGA	AGATAACCAC	GCTTCCATGC	960
CTTGCCTATC	TTGGCATAAT	AGTCTTTGA	CTTGCTAAAT	TTTGAAGGT	CATCCATAAT	1020
CTCTTGTTC	TTGTTGGCT	CCATGGCTAA	AGTATCAAAT	TTGATGGAT	TTCAACAC	1080
TACCTGGCTC	CACATTCTCC	TCCGTATCC	ATACCCACCA	TCTCCCTTAC	TGTTTGTGTA	1140
CAACTTCTC	TGCCCTTC	TTACTGAAAT	TGCCTTCCCT	TCGTCCAATA	CATACTTCAA	1200
GTATGAAGCG	GTGATAAGCT	CGCGGTTCTT	TCTGTGAAAC	TTGAGTTGA	AATACCTCTT	1260
CTCATCCTCC	CTAGGGTACC	AAGAAATTGT	CTGCTGCTG	GCTACTTGT	GGCTAGAAAT	1320
CCACCAGACT	TTCTCGCTT	TATATTCTAC	GGTACCTCC	TCATGATCAT	CCATGGTTAG	1380
TACAAGAGAT	TGGCCATCTT	TCACTACATT	GGCTTGTGAGA	CGCTTAGCTT	GTGTGGAGGA	1440
GTCTTGCTT	AGGTACCTTT	CAATTGCTAC	ATAAGCTT	CTACGCTCGA	ACCAGCCATC	1500
AGTTTCTAC	TCATGAAAAA	TAATGTGCAT	ATAAGGTAG	AAATAGCTCA	CGAGTTTATC	1560
GGTATACCTC	CTAATATGAC	CACGAAGTTC	GTGAGGAAAA	TAGTTCTGGT	ACATGGTCCA	1620
GGCAAACATG	ATTGTTGCAA	TAGCTGGACC	CAAATGAGTC	CAAACATCTT	GCATCATCAT	1680
CATCATCTAA	TTTCTCT					1697

(2) INFORMATION FOR SEQ ID NO: 15

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 654 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- 63 -

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TTAACCTTGA GATAAGCATT	AAAAAAACTC AATGGCAGGG	AAGGTGAGA AAGTGCTTGC	60			
AGTACTGATG CTTGCAATGC	TTCTGTTTC	GGAGCATTIC ATGGCTGCTA	ATCATGAAAT	120		
TAAAACAAC	GAAGATAACT	CTATTAGCCC	TTTCCTGCITA	ATAAAATGTT	TATATGGATG	180
CAGGGGTTG	CCACCTGCAA	AAGCAGCCAT	TTGTGCAGCT	CAATGTTTGT	TTAAGTGC	240
TGTCCAAGAT	GAGGCCAATA	TAGCTGAAAC	TAAGGGCATA	ATAGGTGAGA	CTGCATACAA	300
CCAGTATGAT	GTTGGATGTG	CCCTTGGCTA	CTGCTCTGAG	TTCCCTGTTGA	ATTATGATGA	360
GAGGAGGTTC	AACTGCTGCA	TGGAATACTG	TCGCGAGGGC	AAAATGACCT	GTCCTGTTGA	420
GGCTGCACCT	TGAAGAAATG	GTTGCCCTAA	AATTATGCC	TCATCAAATG	GAAGTACACT	480
GCTTTTCTA	CTTCGGTGT	TTAGTAGTAG	TAGTAAATAA	GTGAGGCATG	TTACGTACTC	540
TTATGTTTG	TAATAATTAT	GCTTTTTAAT	AATGTAATCT	GTCTGTGTC	ATACAATGCA	600
CACGACGCTA	GCTACTACTT	TTTATCTACT	AAAAACGAAA	AGTAACTTAT	TTCT	654

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1031 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GAGACAAATA	CCTATAATAA	GCATTCTATA	ATCTTCTTGC	TTCTTGTCA	GAATAATGGG	60
GAATCTTTTC	TGTTGCGTGC	TTGTGAAGCA	ATCAGATGTT	GCGGTCAAGG	AGAGATTGG	120
CAAATTCCAA	AAAGTACTTA	ATCCAGGTCT	CCAATTGTT	CCATGGGTCA	TCGGTGATTA	180
CGTCGCCGGT	ACACTGACCC	TTCGTCCTCA	GCAACTCGAT	GTTCAGTGTG	AAACCAAAAC	240
AAAGGACAAT	GTGTTGTGA	CAGTGGTGC	ATCCATACAA	TACAGAGTCT	TAGCTGACAA	300
GGCAAGTGAT	GCTTTTACA	GACTCAGCAA	TCCAACCAC	CAAATCAAAG	CCTACGTCTT	360

- 64 -

TGATGTGATC AGAGCATGTG TTCCAAAGCT GAACTGGAC GATGTGTTG AGCAGAAGAA	420
TGAAATTGCC AAACTGTGG AAGAAGAGCT AGACAAAGCC ATGACTGCTT ATGGTTACGA	480
AATCCTCAA ACCCTAATTA TCGACATTGA GCCTGATCAA CAGGTAAAC GTGCCATGAA	540
CGAAATCAAC GCGCGGCGA GGATGAGAGT GGCAGCGAGC GAAAAAGCAG AGGCTGAGAA	600
AATCATTCAG ATCAAAAGAG CAGAGGGTGA AGCAGAGTCA AAGTACCTGT CGGGACTCGG	660
AATCGCTCGG CAGAGACAAG CGATCGTGGA CGGTCTTGAG AGACAGTGT CTTGGGTTCG	720
CAGGAAACGT GCCAGGGACG TCAGCGAAGG ATGTGTTGGA CATGGTGAATG ATGACTCAGT	780
ACTTTGACAC AATGAGAGAT ATCGGAGCAA CTTCTAAATC CTCGCGGTG TTTATCCCTC	840
ACGGTCCAGG CGCCGTCTCT GACGTGGCAG CGCAGATTG AAATGGATTA TTACAGGCCA	900
ACAATGCCTC CTAATCACTC AAGTCAAATT GTCTGGTGTG TCTCTTTATA TATTTTCGTA	960
TCTTCTTATT AAAAAGGTAA ATTTGACTTT TAATATAATG GTGTGCCTAT TGGGAAAAAA	1020
AAAAAAAAAA A	1031

## (2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 860 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GGGAACAAAA GCTGGAGCTC CACCGCGGTG GCGCGCTCT AGAACTAGTG GATCCCCCGG	60
GCTGCAGGAA TTGGCAGCA GCAACTTAGA AAAATGAATT TTACTGGCTA TTCTCGATT	120
TTAACGCTCT TTGTAGCTCT TGTAGGTGCT CTTGTTCTTC CCTCGAAAGC TCAAGATAGC	180
CCACAAGATT ATCTAAGGGT TCACAAACCAAG GCACGAGGAG CGGTAGGCGT AGGTCCCATG	240
CAGTGGGACG AGAGGGTGTG AGCCTATGCT CGGAGCTACG CAGAACAACT AAGAGGCAAC	300
TGGAGACTCA TACACTCTGG TGGCCCTTAC GGGGAAACT TAGCCTGGGG TAGCGGTGAC	360
TTGTCGGCG TCTCCGCGT GAACATGTGG GTTAGCGAGA AGGCTAACTA CAACTACGCT	420
GCGAACACGT GCAATGGAGT TTGTGGTCAC TACACTCAAG TTGTTGGAG AAAGTCAGTG	480
AGACTCGGAT GTGCAAAGT GAGGTGTAAAC AATGGTGGAA CCATAATCAG TTGCAACTAT	540

- 65 -

GATCCTCGTG GGAATTATGT GAACGAGAAG CCATACTAAT GAAGTAATGA TGTGATCATG	600
CATACACACG TACATAAAGG ACGTGTATAT GTATCAGTAT TTCAATAAGG AGCATCATAT	660
GCAGGAYGTA TCAATATTTA TCAAATAATA CAAATAAGAG CTGAGATTAC GAGAATCTAT	720
TTAAATTAAA AGTTACATAC TTAATTATTA TAGTTATATA TGAAAATAT GTGGCTTTT	780
TAAAAGTTAC ATAATTAAATT ATTATAGTTA ATGTCTTCA AAAAAAAA AAAAAAAACT	840
CGAGGGGGGG CCCGGTACCC	860

## (2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: oligonucleotide primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

TTTGGCGTCT TCCATTTTC TAAGTTGATA ATGG

## (2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: oligonucleotide primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TATCAACTTA GAAAAATGGA AGACGCCAAA

## (2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1064 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- 66 -

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAACAATGAA	CATTAAGGTA	TCATTACITTT	TCATTCTTACCA	AATATTCTTG	CTTCTCTAA	60
CCAGCAAAGT	AAAAGCTGGA	GATATTGTAG	TCTATTGGGG	CCAAGATGTA	GGAGAAGGTA	120
AATTGATTGA	CACATGCAAC	TCTGGCTCT	ACAATATTGT	CAACATTGCC	TTTTTATCTT	180
CCTTGGCAA	TTTCCAAACT	CCTAAACITTA	ACITAGCTGG	CCATTGTGAA	CCATCTCTG	240
GTGGITGCCA	ACAGITGACA	AAAAGCATCA	GACATTGTCA	AAGCATAGGC	ATTAAAATCA	300
TGCTCTCCAT	TGGAGGTGGA	ACTCCTACCT	ACACATTATC	CTCAGTTGAT	GATGCCAGAC	360
AAGITGCTGA	TTACCTGTGG	AAACAATTTC	TCGGCGGCCA	ATCATCTTTT	AGGCCACTTG	420
GAGATGCTGT	ATTAGATGGC	ATAGATTTC	ATATTGAAC	TGGCCAACCA	CATTATATTG	480
CACTGCCAG	GAGACTTTCA	GAACATGGCC	AACAAGGTA	AAAATTATAC	TTAACTGCAG	540
CACCAACAATG	TCCTTTCT	GATAAACITTC	TTAATGGTC	ATTGCAAAC	GGTTTATTG	600
ACTATGTTG	GGTCCAATT	TACAACAATC	COGAGTGCAG	GTTCATGAGC	AATTCAAGAAA	660
ATTCAAGAG	GAGGTGGAAT	CAGTGGACAT	CAATCCCTGC	AAAGAAGTGT	TATATTGGAC	720
TTCCAGCAGC	CAAGACAGCC	GCGGGTAATG	GCTATATTCC	AAAGCAAGTG	CTAATGTCAC	780
AAGTTTACC	ATTCTAAAG	GGGTCTCAA	AATATGGAGG	TGTCATGCTT	TGGAATAGAA	840
AATTGATGT	CCAATGTGGC	TATAGCTCTG	CTATCAGGGG	TGCTGTTAA	GTTCTGAATG	900
AACAAGGCGC	CCCTGAATCG	CTATAAGCCA	TCGTTAAGGC	CTAAATAAAG	CAAGTTAATT	960
TGCTGTTATC	TGCCTAGAAA	GTACTTAAGT	TTTAATTGT	ACTGATGAAA	ATGTGAAGGT	1020
CATCTTGTTT	CCITCTGAT	AAATAGTAGTA	CTATGGTTCT	CITT		1064

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1018 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

- 67 -

TGAAGTGT	TA GCACACACAC	ACAAAAAAA	CITAAAGTTA	TGATCAAATA	TAGTTTCTT	60
TTGACAGCAT	TAGTGCTATT	TCTTCGAGCA	TTAAAAC	TAG AAGCAGGGGA	TATAGTAATA	120
TATTGGGCC	AAAATGGAA	TGAAGGTAGC	TTAGCTGACA	CTTGTGCAAC	AAATAACTAT	180
GCCATTGTCA	ATATTCGTTT	CCTTGAGTT	TTTGGGAATG	GCCAAAATCC	AGTGCTAAAT	240
TTAGCTGGTC	ATTGTGATCC	AAATGCTGGT	GCATGCACTG	GCTTAAGCAA	TGACATTAGA	300
GCTGTCAAA	ACCAAGGCAT	CAAAGTTATG	CTTCTCTTG	GTGGTGGTGC	TGGAAGCTAT	360
TTTCTTCTT	CTGCTGATGA	TGCTAGGAAT	GTGGCAAATT	ATTTGTGGAA	CAATTATCTT	420
GGAGGTCAAT	CAAACACACG	TCCACTAGGA	GATGCGAGTC	TAGATGGAAT	TGATTTTGAT	480
ATAGAAGGCG	GGACAACACA	ACATTGGGAT	GAATTAGCAA	AAACTCTATC	ACAATTAGC	540
CAACAAAGGA	AAGTATACTT	AACTGCGAGCT	CCACAAATGTC	CATTCCCAGA	TACATGGTTA	600
AATGGGGC	ACTTCCACTGG	CTTATTGAT	TATGTTGGG	TICAATTITTA	CAATAATCCA	660
CCGTGTCAAT	ACTCCGGTGG	GAGCGCGGAC	AATTAAAAAA	ATTACTGGAA	TCAGTGGAAC	720
CCGATTCAAG	CTGGAAAAAT	TTTCTGGGA	TTGCCAGCAG	CTCAAGGAGC	AGCTGGAAGT	780
GGTTTTATAC	CATCTGATGT	TCTTGTTCT	CAGGTTTAC	CATTAATTAA	TGGTTCACCA	840
AAGTATGGGG	GTGTTATGCT	TTGGTCTAAA	TTTTATGACA	ATGGTTATAG	CCTGTCTATT	900
AAGGCTAATG	TTTGAGATAT	ATGATCATAG	CTAGTCAGCT	TGTATTAATA	TGATGACGTC	960
AATAATGT	TA TATAAAC	TATATAGTAC	TCAATAATAA	GGCTTGAAA	GTTACTTA	1018

## (2) INFORMATION FOR SEQ ID NO: 22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

GGGATCCCTG CA

12

## (2) INFORMATION FOR SEQ ID NO: 23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs

- 68 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

18

ATAGTCCTGT TGAGAGTT

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

20

TCACGGGTG GGGTTCTAC

(2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

18

AGGAGATGGT TTGGTGGA

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- 69 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ATACGTTCTA CTATCATAGT

51

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

His Phe Cys Tyr Ile Glu Glu  
1 5

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Ile Arg Ala Ile Asn Gly  
1 5

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 731 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 40..528

- 70 -

## (D) OTHER INFORMATION:/note= "PR-1mz cDNA clone"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GAATTCCGGG AACTAGTAAT TATCAGCAAA CAACAAACAA TGGCACCGAG GCTAGCGTGC	60
CTCCTAGCTC TGGCCATGGC AGCCATCGTC GTGGCGCGT GCACGGCCCA GAACTCGCCG	120
CAGGACTACG TGGACCCGCA CAACCGGGCG CGCGCGGACG TGGGCGTCGG GCCGGTGTCC	180
TGGGACGACA CCGTGGCCGC GTACCGCGAG AGCTACGCGG CGCAGCGCCA GGGGCGACTGC	240
AAGCTGATCC ACTCCGGCGG GCGCTACGGC GAGAACCTCT TCTGGGGCTC CGCOGGGCGCC	300
GACTGGTCGG CGTCCGACGC CGTGGGCTCC TGGGTTGTCGG AGAAGCAGTA CTACGACAC	360
GACACCAACA GCTGCGCGGA GGGGCAGGTG TGCGGCCACT ACACGCAGGT GGTGTGGCGC	420
GACTCCACCG CCATCGGCTG TGCGCGCGTC GTCTGCGACA ACAACGCGG CGTCITCATC	480
ATCTGCAGCT ACAACCCGCC GGGCAACGTC GTCGCGAGA GCCCCTACTA GACTGTTCATG	540
CATACTACAA TTATATATAT ATACGCTTAA TATTAACITTC AGCATGCATG CATATTATAA	600
ATAGTGTGT CAACTCTGTA TATCATTATA TTACGATGAT TATATTGTG ATGAACATAA	660
TAAGTCAATA AAACCATGCA GGCGATTGT GAGCCAAAAA AAAAAAAA AAAAAAAA	720
AAAAACTCGA G	731

## (2) INFORMATION FOR SEQ ID NO: 30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Met Ala Pro Arg Leu Ala Cys Leu Leu Ala Leu Ala Met Ala Ala Ile			
1	5	10	15
Val Val Ala Pro Cys Thr Ala Gln Asn Ser Pro Gln Asp Tyr Val Asp			
20	25	30	
Pro His Asn Ala Ala Arg Ala Asp Val Gly Val Gly Pro Val Ser Trp			
35	40	45	
Asp Asp Thr Val Ala Ala Tyr Ala Gln Ser Tyr Ala Ala Gln Arg Gln			
50	55	60	
Gly Asp Cys Lys Leu Ile His Ser Gly Gly Pro Tyr Gly Glu Asn Leu			

- 71 -

65	70	75	80
Phe Trp Gly Ser Ala Gly Ala Asp Trp Ser Ala Ser Asp Ala Val Gly			
85		90	95
Ser Trp Val Ser Glu Lys Gln Tyr Tyr Asp His Asp Thr Asn Ser Cys			
100	105		110
Ala Glu Gly Gln Val Cys Gly His Tyr Thr Gln Val Val Trp Arg Asp			
115	120	125	
Ser Thr Ala Ile Gly Cys Ala Arg Val Val Cys Asp Asn Asn Ala Gly			
130	135	140	
Val Phe Ile Ile Cys Ser Tyr Asn Pro Pro Gly Asn Val Val Gly Glu			
145	150	155	160
Ser Pro Tyr			

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 831 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- iii) HYPOTHETICAL: NO
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 72..587
  - (D) OTHER INFORMATION:/note= "PR-5mz cDNA clone"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

GAATTCCCGGA	GCCTAGCTCT	ATAGCTCGAG	TATTGCTTGC	TCTCCACAAG	AAACAAGGCA	60
GAGCACCAAC	AATGGCCGCC	CGTCCTCGG	TCCTCCTGCT	GCTCCTGGCC	GCGCCCTTGG	120
CTGGCATGAG	CGCCAACGCC	GCCACCTTCA	CCATCACCAA	CAACTGCGGG	TTCACCGTGT	180
GGCCCGCGGC	CACCCCGGTG	GGCGGGGGCA	CGCAGCTGAA	COCGGGGGGG	ACGTGGACCG	240
TCAACGTGOC	GGCOGGCACC	AGCTCOGGCC	GCGTGTGGGG	COGCACCGGC	TGCTCCCTICA	300
ACGGCAACAG	CGGGAGCTGC	CAGACGGGCG	ACTGCGGGGG	CGCGCTGGCC	TGCACGCTGT	360
CGGGGCAGOC	GCGCCTGACG	CTGGCGAGT	TCACCCATCGG	CGGCAGCCAG	GACTTCTAOG	420

- 72 -

ACATCTCGGT CATCGACGGC TACAACCTCG CCATGGCCTT CTCCCTGCAGC ACCGGCGTGC	480
GGCTGGTGTG CACGGACCCG GGATGCCCG ACGGGTACCA CAACCCCCCC GACATGAAGA	540
CCCATGCCTG TGGGGGCAAC AGCAACTACC AAGTCACCTT CTGCCCCGTGA TCGGCCGGAA	600
TAGGCTCTGC ATGCATGGGC TAGTGATTIG GATTTTGCAG GAATAATTCT GATTTGGATT	660
TGCACAATAA TATATAAGCG TGCATTACAC GTGCACGCAT TTGTATGTAC GCGCGTATAC	720
GTCGTACCTA CGTAGCTGGA TTTGTGTGCG TGCTTCGTAC CAAATATTAA TAAGAATAAA	780
CAAATATGAC TCGTGTACATA TATATATAAA AAAAAAAA AAAAAACTCGA G	831

## (2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 172 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

Met Ala Ala Ala Ser Ser Val Leu Leu Leu Leu Leu Ala Ala Ala Leu	
1 5 10 15	
Ala Gly Met Ser Ala Asn Ala Ala Thr Phe Thr Ile Thr Asn Asn Cys	
20 25 30	
Gly Phe Thr Val Trp Pro Ala Ala Thr Pro Val Gly Gly Gly Thr Gln	
35 40 45	
Leu Asn Pro Gly Gly Thr Trp Thr Val Asn Val Pro Ala Gly Thr Ser	
50 55 60	
Ser Gly Arg Val Trp Gly Arg Thr Gly Cys Ser Phe Asn Gly Asn Ser	
65 70 75 80	
Gly Ser Cys Gln Thr Gly Asp Cys Gly Gly Ala Leu Ala Cys Thr Leu	
85 90 95	
Ser Gly Gln Pro Pro Leu Thr Leu Ala Glu Phe Thr Ile Gly Gly Ser	
100 105 110	
Gln Asp Phe Tyr Asp Ile Ser Val Ile Asp Gly Tyr Asn Leu Ala Met	
115 120 125	
Ala Phe Ser Cys Ser Thr Gly Val Arg Leu Val Cys Thr Asp Pro Gly	
130 135 140	
Cys Pro Asp Ala Tyr His Asn Pro Pro Asp Met Lys Thr His Ala Cys	
145 150 155 160	

- 73 -

Gly Gly Asn Ser Asn Tyr Gln Val Thr Phe Cys Pro  
 165 170

## (2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 790 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..790
- (D) OTHER INFORMATION:/note= "cDNA called PSI-1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

GGCACGAGGT CATATTCGTT GAAAACATCA TTCACTCTTC TCTGTGTCTC CTTCTTTCAC	60
CCGAAATTGT CTCTGACACC GNCTTGTNC AGNCATGGCT AATGCTGCGT CAGGAATGGC	120
AGTCCATGAT GACTGCWRY KARKAKYYCY KGGAMTGAAG GCGAAAAGGA CACAMCRYTT	180
CATTGCTAC AAGATTGAGG AGAACGAGAA GCAAGTGATT GTTGAGAAAG TTGGTGAACC	240
TATTCTAACT TACGAGGACT TTGCAGCAAG TCTTCCAGCT GACGAATGMC GATAGCCAT	300
YTATGATTTC GACTTTGTCA CTGCAGAGAA TTGCCAGAAG AGCAAGATT TCTTCATTGC	360
ATGGTGTCCC GACGTAGCAA AGGTGAGAAG CAAGATGATC TATGCGAGCT CTAAGGACAG	420
GTTCAGCGT GAACCTGATG GAATTCAAGT GGAGCTTCAA GCAACTGATC CAACTGAGAT	480
GGATCTTGTAT GTTTGAAAAA GCGCGTCAA CTAAAAGAAA CTCTCTTGA ATAAGCTTCC	540
GATTATATTTC GACTATTATG AAATGGTGT TTCTTATTCTC TTACAGTCTC TTGTGACTGT	600
TAGATACACT TGACACCCCTT CATTCTATCT GTCTGGTCTG TTCTCTGTGT CTATGTTACC	660
TTACAGTGTGTC TCATCTATAT TCAAAATTAT CACAAGGTTT GTGCTGTGTT TGCTTCCCTT	720
GTTCAAACTC TACCAAAGCT TAAGATTAAG ATTTGGTAA GATGATTCTT TTAAAAAAA	780
AAAAAAAAAA	790

## (2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 base pairs
- (B) TYPE: nucleic acid

- 74 -

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:1..508
- (D) OTHER INFORMATION:/note= "cDNA called PSI-2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

CCTCCAAGAA CCTAGTGGGA TCCCCCGGGC CTGCAGGAAT TCGGCACCGA GCTCCTGCC	60
GAAATTCCGC ACGAGTGTG ATCTAAATCC TCACTAGGTG GTTCCGTTTA AGCTTAGCAA	120
CGCCAAAGAT CTAGACAATA ATTGCAAGTT ACCGCCTTGA TCCATCTTC CGTCCCGAAG	180
CCGTCCCTGA TCCAGCCCAA GATGTTTGTG GGTAGATGA CAAAGTAGAT CCACGGGTAT	240
ATATTATGGT AAAGTTGGCG GCGAAGGGAT TGAAGTTCAA TAATGATATG TTTTCCTGGTG	300
GTTGCATCCC ATCTGAAGTG GTTATAGCCC CCAAAACAA AAAATCTCTA CAATTTGAA	360
AGGTGTTCGT GGCAGGAAAT CTGCTAAGCA AGGAAGGATC TCCAAGAAC CAACTGGGAA	420
TAGAAGGATC CATCAATCAT TCTCTAAATGA GGAGGAATCT GTTTAACTTT GTTATTGACT	480
AATAATCCTA TTATCGCCTA AAGCTGTT	508

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 570 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION:1..570
- (D) OTHER INFORMATION:/note= "cDNA called PSI-3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GGCACGAGCT CGTGCCTGTC GTGCCGATT CGGCACGAGG CAACATTGCAACACCATCG	60
ACGGTGATAG GCCTCGGAGG ATCAATCCATC ACCACCAAAC CCTTCTCTTC ATCCCTTTTA	120

- 75 -

AAACCAACAT TAAGCGCCAA GAACCCTTTG AGACTCGCCG GTGCATCGGG AGGAAGAGTC	180
ACTTGCCTTG AGAGGAAC TGTTGAGGAGA GATTGAAACG TGGTAGGATT TGGGCTGATC	240
GGATGGCTAG CTCCGTCGAG CATTCCAGCG ATAAATGGGA AGAGCCTGAC GGGTCCTTC	300
TTCGATAGCA TCGGAAC TGCA GCTCGCTCAC TTCCCGACTC CTCCAGCTCT CACITCACAG	360
TTCTGGTTGT GGTTGGTTAC GTGGCCTTAA GGCCTCTTC TCTGCCTCAC TTTGGACAA	420
ATCGGATTCA AGGGCAGGAC TGAAGATTAC TTCTAAGGAT AACTATTCTT GTTTTCGTTT	480
GTACTATATG CTCTCTCTGG KTATGTGAA TATTATCAAT CAAAACAAAG CTTTCCTTGC	540
TATTTGATGT TTCATTCTAT AAAAAAAA	570

## (2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1418 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1..1418
  - (D) OTHER INFORMATION:/note= "cDNA called PSI-4"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GGCACGAGCA GAATCCCTAG CATCAGAGAA TTCTTCAATG ACAGACACTT ATAACCAGCA	60
GAGAGGTCTT GTCAACCAAC TAAAAGATGA CATGGAGAGG CTATATCAAC AGATCCAAGC	120
CCAAATGGGC GAACTTGAGT CTGTCAGGGT TGAGTATGCA AACGCGCAGC TAGAATGTAA	180
TGCTGCTGAT GAGCGTCCC AAATACTGGC TTCTGAAGTC ATCAGTTGG AAGATAAGGC	240
TCTCAGACTC AGGATCTAAT GAGTTAAAGC TGGAGAGGGA ACTGGAGAAG GCACAAAACA	300
GAATTGTTA TCTTACAAGA AAAAATTGCA GGACCTTAA GAAAGGATCC GTCAAAGATT	360
TACAAATCTT ACTATTAAG CTCTTCAGGA AAGAGAAAGG AAGGGTCCCT ACAAAACTAT	420
GGGTGCAGGA AAGCTTCATC TGGTGGGAA AAATCCACTG GATCTTGTAG TIAAAACTCA	480
ACTAGCAGAA AGAACGTGTC AACCTCTACA GAAGGTCTTG CAATCTCAGA TACTACGCCA	540
GAGAGCTCCA ACCAGGAAAC AGATTCTACT ACTCTGCTCG AAAGTGATTC ATCTIWATACA	600

- 76 -

GCTATCATTG	CTGAAACTAG	ACAAATTAAC	CITGAAGGCT	TTTCATTGAG	CGTCCCAGCT	660
GATCAGATGA	GAGTGATTCA	TAACATTAAT	ACGCTGATTG	CTGAGTTGGC	AATTGAGAAG	720
GAAGAACTGG	TGCAAGCACT	GTCATCTGAG	TTATCTGAA	GTGCGCATGT	GCAGGGAGCTG	780
AACAAAGAGT	TATCCAGAAA	ACTTGAAGCG	CAGACGAAA	GGTTAGAGCT	TGTAACAGCG	840
CAGAAGATGG	CCATAGACAA	TGTTTCACCA	AAAAAGCAGC	AGSCGGACAC	TCATGTTGTT	900
CAAGAGAGAA	CMCCGATTGC	AGATGAAGGC	GATGAGGTGG	TAGAAAGGGT	TCTWAGGATG	960
GATCATGAAG	ATGTTCCAG	GAGGACOGTC	AAAAAGAAGG	ACAAGCAAGC	TTCTCTAATG	1020
CATGGATGGC	TTGCTCAGTC	GTGACCCCTG	TTGCCATTCT	GTATATCCCG	GGCACAGCCA	1080
TGGATTGCTT	CTTCTCCTTT	AGKKKTGGAA	GAATTCAATT	GGTTCTTTG	ATCCCCAAAT	1140
TTTGGATTTC	ATTGGATTGA	GOMCAAATGG	TATAGTGAGA	AACACCATGA	ACTTGAAGAG	1200
TTCCAAATGG	ATAGAGGAGA	CTCGGAAGCT	AAACAGAGA	CGAAGGTAGT	TGTTATATGA	1260
GTGOGTTTTT	GTCTGAATT	TGTAATTGTT	GAATATATAG	CCAAGTTAAA	GTGTTCCCTCC	1320
TAGCAAATAT	AATTTACCAA	AAAAAAATCT	TGTATGGTTT	CGGATATATT	TTATCTCTC	1380
GTTTTAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAA	AAAAAAA		1418

## (2) INFORMATION FOR SEQ ID NO: 37:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 913 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION: 1..913
- (D) OTHER INFORMATION:/note= "cDNA called PSI-5"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

GGCACGAGAG	ACAATCAGAT	TCTTTACTCT	TACAACGGTG	GAGACCAAAC	CAACGAGAGT	60
CTTGCTGCTC	TGTGTTAGA	AAAGTCTCCT	CCTTTTCATA	CTTGGTACTT	TGAAACIATT	120
GGTAAAAGAA	GATTTGGGTT	TTTAATTGGA	GATGGGTTTG	TTTATTTCCG	CATTGTTGAT	180
GAGGTTTIGA	AGAGATCTAG	TGTTCTTAAG	TTCTTGAAC	ATTTGGAGAG	ATGAATTAA	240

- 77 -

GAAAGCTGCT AGGGAGAATT CTAGAGGAAG TTTCACTGCT ATGATTGGTT CTATTAATGT	300
TGAAGATCAG CTTGTTCTTG TTGTTACTAG ACTTATAGCT TCTCTTGAAC GTGTTGCTGA	360
GAGTAGTAGT AACAAATGAAT TGAAGAGTAG TAATCTTGGT GAACAAAGCG AAGGTTGAA	420
TTCTACTAAA GCTCCGTTGT TGGGGAGGTT AAGTAAGCAA GAGAAGAAGA AAGGGAAAGA	480
TCATGTGATT GAGCTTGAGG AACACAGGAA GTCIAATGAT AGAGGGAAACA TAACCGATGA	540
TTCAGCGGGA GCTGGAACGT CGTTGGAGAA GGAGTGTGTG TCGAGTAGTG GACGTTCTGT	600
TACTCAAAGC TTTGAATGGA AATGGCGGCG GYTAGTTCAAG ATTGTTCTTG CTATTGATGC	660
AGCTATTGTC TTGACACTGT TTGGTATTTG GCTGGCTATA TGTGGGGTA TCGAGTGTAC	720
ACGTTCGTGA TTGTGCAGAC AGAAAATCGG ATAAACAGAG TCGATTCTTT ACATTCCCTT	780
GTGCATTCCA ATAGCCAAA AATATGTTGA TTGTAATGTG AACCATTCCTT AAGTTATAAT	840
TTGCCCTTATG TATTGCTTTA GTCACAGCTT TTGAAATATG TACAGTCGGA TTCTAGAAAA	900
AAAAAAAAAA AAA	913

WHAT IS CLAIMED IS:

1. A recombinant or chimaeric DNA molecule comprising a plant SAR gene different from the DNA molecule p1.4.3 (SEQ ID NO. 12 / SEQ ID NO. 13), wherein the wild-type gene corresponding to said SAR gene can be chemically induced in a plant in a protein-synthesis independent manner.
2. A DNA molecule according to claim 1, wherein the corresponding wild-type gene can be chemically induced in a plant despite the presence of the protein synthesis inhibitor cycloheximide.
3. A DNA molecule according to claim 1, wherein the corresponding wild-type gene can be induced in a plant by salicylic acid, isonicotinic acid or methyl benzo-1,2,3-thiadiazole-7-carboxylate.
4. A cDNA molecule according to claim 1.
5. A DNA molecule according to claim 1 comprising DNA selected from the group of DNA molecules p11.30.13 (SEQ ID NO. 11), p66B1 (SEQ ID NO. 14), pDPA2 (SEQ ID NO. 16), PSI-1 (SEQ ID NO. 34), PSI-2 (SEQ ID NO. 35), PSI-3 (SEQ ID NO. 36), PSI-4 (SEQ ID NO. 37), and PSI-5 (SEQ ID NO. 38).
6. A recombinant or chimaeric DNA molecule comprising a plant SAR gene different from the DNA molecule p1.4.3 (SEQ ID NO. 12 / SEQ ID NO. 13) obtainable by a method comprising
  - (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;
  - (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
  - (c) identifying clones which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.

7. A recombinant or chimaeric DNA molecule comprising a plant SAR gene different from the DNA molecule p1.4.3 (SEQ ID NO. 12 / SEQ ID NO. 13) obtainable by a method comprising
  - (a) cloning genes which are chemically inducible in the presence or absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;
  - (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
  - (c) identifying amplification products which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.
8. A recombinant or chimaeric DNA molecule comprising a plant SAR gene, wherein the wild-type gene corresponding to said SAR gene can be chemically induced in a plant in a protein-synthesis dependent manner.
9. A DNA molecule according to claim 8 wherein the corresponding wild-type gene can only be chemically induced in the absence of the protein synthesis inhibitor cycloheximide.
10. A DNA molecule according to claim 8 wherein the corresponding wild-type gene can be induced by salicylic acid, isonicotinic acid or methyl benzo-1,2,3-thiadiazole-7-carboxylate.
11. A cDNA molecule according to claims 8.
12. A DNA molecule according to claim 8 comprising DNA selected from the group of DNA molecules p1.1.1 (SEQ ID NO. 9), p.11.31.4 (SEQ ID NO. 10), and p.14.22.3 (SEQ ID NO. 15).
13. A recombinant or chimaeric DNA molecule comprising a plant SAR gene obtainable by a method comprising
  - (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;

- (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying clones which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

14. A recombinant or chimaeric DNA molecule comprising a plant SAR gene obtainable by a method comprising

- (a) cloning genes which are chemically inducible in the absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;
- (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying amplification products which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

15. A recombinant or chimaeric DNA molecule comprising the cDNA of a wild-type wheat gene which wild-type gene can be chemically induced in a plant.

16. A DNA molecule according to claim 15 wherein the corresponding wild-type gene can be induced in a plant by methyl benzo-1,2,3-thiadiazole-7-carboxylate.

17. A DNA molecule according to claim 15 comprising DNA selected from the group of DNA molecules WCI-1 (SEQ ID NO. 3), WCI-2 (SEQ ID NO. 4 / SEQ ID NO. 5), WCI-3 (SEQ ID NO. 6), WCI-4 (SEQ ID NO. 7), and WCI-5 (SEQ ID NO. 8).

18. A recombinant or chimaeric DNA molecule encoding *Arabidopsis* class IV chitinase, *Tobacco* chitinase/lysozyme, *Maize* PR-1, or *Maize* PR-5.

19. A DNA molecule according to claim 18 encoding *Arabidopsis* class IV chitinase type A or B.

20. A DNA molecule according to claim 18 comprising the DNA sequence pChit4-TA (SEQ ID NO. 1), pChit4-TB (SEQ ID NO. 2), pBSCL2 (SEQ ID NO. 20), or pBSTCL226 (SEQ ID NO. 21).
21. A DNA molecule according to claim 18 comprising the DNA sequence PR-1mz (SEQ ID NO. 30), or PR-5mz (SEQ ID NO. 32).
22. A recombinant or chimaeric DNA molecule comprising the promoter region of the *Arabidopsis* PR-1 gene.
23. A DNA molecule according to claim 18 comprising 4.2 kb upstream of the start codon of the PR-1 wild-type gene.
24. A method for obtaining a recombinant or chimaeric DNA molecule encoding a gene which corresponds to a wild-type gene which can be chemically induced in a protein-synthesis independent manner, comprising
  - (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;
  - (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
  - (c) identifying clones which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.
25. A method for obtaining a recombinant or chimaeric DNA molecule encoding a gene which corresponds to a wild-type gene which can be chemically induced in a protein-synthesis independent manner, comprising
  - (a) cloning genes which are chemically inducible in the presence or absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;
  - (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and

- (c) identifying amplification products which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.

26. A method according to claims 24 or 25, wherein cycloheximide is used as the protein synthesis inhibitor.

27. A method for obtaining a recombinant or chimaeric DNA molecule encoding a gene which corresponds to a wild-type gene which can be chemically induced in a protein-synthesis dependent manner, comprising

- (a) cloning genes which are chemically inducible by differential screening of cDNA libraries;
- (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying clones which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

28. A method for obtaining a recombinant or chimaeric DNA molecule encoding a gene which corresponds to a wild-type gene which can be chemically induced in a protein-synthesis dependent manner, comprising

- (a) cloning genes which are chemically inducible in the absence of a protein synthesis inhibitor by differential display of isolated RNA using polymerase chain reaction;
- (b) further analyzing the amplification products obtained in step (a) by Northern hybridisation to RNA isolated from cells chemically induced in the presence or absence of a protein synthesis inhibitor; and
- (c) identifying amplification products which hybridise with RNA chemically induced in the absence of a protein synthesis inhibitor but not with RNA chemically induced in the presence of a protein synthesis inhibitor.

29. A method according to claims 27 or 28, wherein cycloheximide is used as the protein synthesis inhibitor.

30. A method for obtaining a recombinant or chimaeric DNA molecule comprising the cDNA of a wild-type wheat gene which wild-type gene can be chemically induced in a wheat plant, comprising

- (a) cloning genes which are chemically inducible by differential screening of wheat cDNA libraries;
- (b) further analyzing the clones obtained in step (a) by Northern hybridisation to RNA isolated from wheat cells chemically induced; and
- (c) identifying amplification products which hybridise with both RNA chemically induced in the absence of a protein synthesis inhibitor and RNA chemically induced in the presence of a protein synthesis inhibitor.

31. A method for obtaining a recombinant or chimaeric DNA molecule encoding plant class IV chitinase comprising amplifying by the polymerase chain reaction DNA molecules from plant DNA using 5' and 3' oligonucleotides designed degenerate for the peptides HFCYIEE (SEQ ID NO. 27) and IRAING (SEQ ID NO. 28), respectively.

32. A method for obtaining a recombinant or chimaeric DNA molecule encoding plant chitinase/lysozyme comprising screening a plant cDNA library with a cDNA probe encoding cucumber chitinase/lysozyme.

33. A method for obtaining a recombinant or chimaeric DNA molecule comprising the promoter region of the *Arabidopsis PR-1* gene comprising

- (a) screening of an *Arabidopsis* cDNA library with a probe of *Arabidopsis PR-1* c-DNA; and
- (b) identifying a clone carrying the promoter fragment by hybridisation to an oligonucleotide corresponding to coding sequences of the *Arabidopsis PR-1* gene.

34. A method of improving protection of a plant against a pest comprising transgenically expressing in said plant two or more DNA molecules encoding anti-pathogenic proteins, wherein the transgenically expressed proteins exert a synergistic effect.

35. A method according to claim 34, wherein two or more DNA molecules according to any one of claims 1-19 are transgenically expressed.

36. A method of protecting plants against pests comprising transgenically expressing in said plant class III chitinase.

37. A method according to claim 36, wherein tobacco plant chitinase III is transgenically expressed.

38. A transgenic plant and the progeny thereof comprising a DNA molecule according to any one of claims 1 to 23 displaying anti-pathogenic properties.
39. A transgenic plant according to claim 38 expressing transgenic Tobacco chitinase/lysozyme.
40. A transgenic plant and the progeny thereof comprising two or more DNA molecules encoding anti-pathogenic proteins, which plant displays synergistic anti-pathogenic properties.
41. A transgenic plant according to claim 40 comprising two or more DNA molecules according to any one of claims 1-23.
42. Seed of a transgenic plant according to any one of claims 38-41.
43. Use of a DNA molecule according to any one of claims 1 to 7 to achieve expression of a protein-synthesis independent plant SAR gene in transgenic plants.
44. Use of a DNA molecule according to any one of claims 8 to 14 to achieve expression of a protein-synthesis dependent plant SAR gene in transgenic plants.
45. Use of a recombinant or chimaeric DNA molecule according to any one of claims 18 to 21 to obtain transgenic plants displaying anti-pathogenic properties.
46. Use of the promoter region according to claim 22 to chemically induce a gene in operable linkage with said promoter.